UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Department of General Psychology

Ph.D. Course in Brain, Mind, and Computer Science

Curriculum: Neuroscience, Technology, and Society

XXXI Series

THE MEMORY SKILLS OF MUSICIANS AND NONMUSICIANS

Supervisor: Prof. Massimo Grassi

Co-Supervisor: Prof. Alessandro Sperduti

Ph.D. candidate: Francesca Talamini

AKNOWLEDGMENT

When I think about the people I want to thank, a long list of names comes to my mind; I will probably forget somebody (so if you don't see your name but you know that somehow you helped me during my three years of PhD, this is the moment I say thank you!).

I want to begin by thanking my supervisor Massimo Grassi, who always believed in me (maybe even more than I did), since we started working together for my master's thesis. Massimo was always supportive, professional and friendly at the same time, and I do not have any doubt that I was really lucky in working with him. He is enthusiastic about many things, and he taught me a lot, in particular how to fight against the "dark side" of science!

A big thank you goes also to Barbara Carretti, another person who was always supportive, with whom I collaborated during these three years. She was my mentor for everything that concerned "memory" and always incredibly human and empathetic. Barbara never hesitated to help me even for small things.

Let's now move to my office in Padua: when I say that I could not have found better colleagues, I am not exaggerating. Therefore, I want to thank Giovanna, Giorgia; Luca, Fabio, and Stefano. From the room 25 to the room 38, we shared amazing moments, they saw me laughing, celebrating, crying and losing motivation, but they were always there helping me, with a hug, or with teaching me something I needed to know.

My BMCS mates: Fabio, Sara and Daniele. We always supported each other, and without them, everything would have been more difficult! In particular, a big thank you goes to Fabio, who is now a good friend, who was there in all the best and worst moments and who is surely part of my succeeding in this experience. I know for sure that I've found a true friend for life!

I want to thank also the people with whom I worked while I was staying in Lyon, in particular, Barbara Tillmann, who allowed me to go to Lyon and live a new incredible experience, working on new things that were really important for my PhD project. She was always really professional and helpful in all the steps of the process.

A special thanks to Anne Caclin, who was there for me every time I needed, she worked a lot for my study and she showed me everything I didn't know with extreme patience and precision. She welcomed me in her équipe and I never felt lost, even when everything was new to me.

Thanks to the PAM members, and CAP members of the CRNL, from whom I could learn a lot, and received a lot of help for everything!

A special thanks also to: Agathe, Salomé, Rémy, Sara, Jérémie, and Caroline for their friendship and support, and for the great help they gave me from the very first day, and later with the subjects' testing. I had some of the funniest moments of my working life with you, and I will always treasure a special memory of all of you.

A big thank you also to all the students who helped me in collecting the data for my experiments in Padua, and all the volunteers that participated to my studies.

At last, but not least, a special thanks to my mom, who always supported me.

I thank you with all my heart, you were a part of my life and if this work exists is also thanks to you!

TABLE OF CONTENTS

ABSTRACT

Musicians seem to have superior abilities than those of nonmusicians, that are not just music-related but that extend to classic auditory and even cognitive tasks, in particular memory tasks. However, concerning memory, results tend to vary depending on the memory system investigated (i.e., long-term, short-term, working memory) and on the category of stimuli that are presented (e.g., verbal, visuospatial).

The present research project investigated the memory skills of musicians and nonmusicians, with the final goal of clarifying which (if there are some) characteristics of musicians are linked to better memory skills and if this advantage is general or depends on specific tasks or content of the tasks.

Study 1 investigated the literature on memory skills of musicians and nonmusicians through a meta-analysis. Three meta-analysis were conducted separately for long-term memory, short-term memory, and working memory. The effect of moderators was also tested; defined as the type of stimuli presented in the memory task (i.e., verbal, visuospatial, and tonal). The three meta-analyses revealed a medium effect-size in working memory and short-term memory (i.e., there is a moderate difference between musicians and nonmusicians) with effect of moderators. The advantage of musicians was larger for tonal and verbal stimuli and smaller for visual ones. In long term memory the effect-size was small, with no effect of moderators.

Study 2 aimed to understand if the advantage found in verbal working memory depended on the modality in which the task was delivered (i.e., stimuli presented auditorily or stimuli presented visually). 18 musicians and 18 nonmusicians performed a digit span task that was presented aurally, visually, or audiovisually. The task was performed with or without a concurrent task (i.e., articulatory suppression). Results showed that musicians had significantly larger spans than nonmusicians regardless of the sensory modality and the concurrent task. Secondary analyses showed that the advantage was more evident when the digits were delivered auditorily and audiovisually.

Study 3 aimed to investigate the individual differences among musicians. In particular, the goal was to understand whether the type of music training (classic vs self-taught) could influence the advantage of musicians over nonmusicians in verbal working memory skills, always taking into account the modality of presentation of the verbal stimuli (i.e., visual vs auditory). 102 young adults participated to the study: 33 reader musicians (i.e., that could read music notation), 33 nonreader musicians (i.e., self-taught, that could not read music notation), and 36 nonmusicians A digit span forward and backward was presented in three different modalities, alike study 2. Results showed that reader musicians, nonreader musicians and nonmusicians performed equally well in the digit span

forward. However, the group interacted with the modality, revealing that reader musicians performed better than nonmusicians in the audiovisual presentation of digits. No other difference was found. In the backward digit span no difference among groups was found.

Study 4 aimed to understand whether the superiority of musicians in short-term memory extends to auditory and visual stimuli that are not verbal and not musical. 36 young adults participated to the study, 24 nonmusicians and 12 professional musicians. A verbal memory task was also included as control measure. In the short-term memory tasks, two sequences of elements were presented, with a short delay in between. The participant had to judge whether the second sequence was the same or different from the first. The types of stimuli composing the sequences where the following: verbal stimuli (i.e., syllables, presented either visually and auditorily); visual contour stimuli (i.e., luminance variations); auditory contour stimuli (i.e., loudness variations); visual nocontour stimuli (i.e., kanji ideograms); auditory nocontour stimuli (i.e., pink noises). Results showed that musicians outperformed nonmusicians in the short-term memory task with the auditory contour and nocontour stimuli, and with the visual contour stimuli.

INTRODUCTION

Learning to play a music instrument requires a lot of energy, commitment, motivation, and constant exercise. It is not surprising that it is only a minority of people who succeed and become a musician. Musicians are therefore experts, and experts can be interesting to study because they can help us understanding how the brain and the mind change after gaining this expertise. For this reason, musicians are more and more studied, with the aim of understanding whether the music training, that leads without any doubts to neuroanatomical changes (e.g., Gaser & Schlaug, 2003), can benefit other-than-musical aspects, such as cognition, emotion, personality and so on (Deutsch, 2013).

Concerning cognition, there are several studies that underlined how musicians have enhanced skills with respect to people who never took music lessons (or who took it minimally, hereafter referred to as nonmusicians). These skills span from memory, to language, spatial abilities, attention and others (e.g., Piro & Ortiz, 2009; Sluming, Brooks, Howard, Downes, & Roberts, 2007 Hansen, Wallentin, & Vuust, 2013,). However, it is not yet clear whether these enhanced skills are a consequence of the music training, if they existed a priori, or if they are consequence of other aspects not directly linked to the music training.

The aim of the current dissertation is to investigate deeply the memory skills of musicians and nonmusicians, through a quantitative and qualitative review of the literature, and through three different empirical studies. In the first chapter there will be an overview on the literature about experts and how they are interesting models in cognitive sciences. Starting from the second chapter and on, I will focus on the studies about music training and musicians, presenting also some theoretical notions, in order to better understand the main focus, which will be about the short-term memory and the working memory skills of musicians. From the fourth chapter on, I will describe the project of my PhD program in its different steps: starting from a metanalysis and then moving to three empirical studies. I will end this dissertation with a general discussion and critic considerations about the present studies and more generally about this field.

CHAPTER 1. EXPERTS

What is expertise? Expertise is a status in which a person is extremely competent in one activity, and this competence is the result of an intensive training. It could be the case of a musician, but not only. The expertise could be either gained after a physical training or a mental/cognitive training, but also both of them. Among the experts frequently investigated in cognitive science and neuroscience we find the already cited musicians, chess players, professional athletes and dancers, and many others. Expertise is, therefore, a consequence of skill learning, that could be either perceptual, motor, and/or cognitive. Once expertise is gained, the individual is able to use appropriate mental representation, strategies, cognitive processing, and/or motor responses, in an automatic and effortless way.

Many researchers investigated differences between experts and non-experts, either at a behavioral level or at an anatomical and functional level (i.e., brain activity and plastic changes). But "when" an individual becomes an expert? Some researchers indicated long periods of intense practice as a "measure" of expertise, that is approximately four hours of daily practice for about 10 years (Ericsson, Krampe, & Tesch-Römer, 1993; Ericsson, 2008). From an anatomical point of view, several studies observed how experts differ from non-experts in gray and white matter density (e.g., in jugglers, Scholz et al., 2009; in golfers Jäncke, 2009; in meditation practitioners, Tang et al., 2012). From a functional point of view, researchers observed that experts tend to have a reduced activation in the brain areas that are used in the first phases of the learning-process. Experts reach an automatism after their training; this lesser use of cognitive resources enables to achieve a higher level of performance. A hypothesis is that thank to this reduction in activation, experts have free resources to carry on the specific task, and therefore they can improve in it. On the contrary, for non-experts, there is a higher demand of controlled processes and attention in order to be able to accomplish the task, processes that are typically involved in the first phases of learning. There is not only a reduction of brain activity observed in experts, but also an increased activity in the areas that are task-specific. This is in line with the hypothesis that automatism can save resources that are needed to carry on the specific task of the expertise domain (Hill & Schneider 2006; Chi, Glaser & Farr 1988).

One of the most famous class of experts investigated in neuroscience, is London's taxi drivers. In a study by Maguire and colleagues (2000), taxi drivers were compared to control subjects who did not drive a taxi. The study showed how taxi drivers had a larger posterior hippocampus, whereas a more anterior region was larger in the control group. Moreover, the volume of the posterior hippocampus was positively correlated with the years of experience of the driver, suggesting that these differences are a consequence of their expertise. The posterior hippocampus is considered to store the spatial representation of the environment, thus it is likely that a long training that requires navigation skills, could change the structure of the brain (Maguire et al., 2000).

This kind of studies which compared experts and non-experts in their anatomical and functional brain differences included other various categories, such as braille readers. It was observed, for example, how the Braille readers have enlarged representation for their index finger, with which they read, in their primary motor area (Pascual-Leone & Torres, 1993). Another class of experts that is often investigated is athletes (see Chang, 2014, for an overview). Structural differences (in comparison to non-athletes) were found in Judo players (Jacini et al., 2009), golfers (Jancke et al., 2009) and mountain climbers (Di Paola, Caltagirone, & Petrosini, 2013). Functional differences between experts and non-experts were reported, for example, also in rugby players (Sekiguchi et al., 2011) racquet ball players (Pearce, Thickbroom, Byrnes, & Mastaglia, 2000) and ballet dancers (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2004).

Musicians, of course, were also largely investigated for assessing brain plasticity after their long (usually around 10 years), complex (it requires to integrate multisensory information and specific motor actions), and intense (usually there is an intense daily practice) training (for a review, see Habib & Besson, 2009; Herholz & Zatorre, 2012; Jäncke, 2009). Several studies considered the anatomical differences between adult musicians and nonmusicians, but also shorter periods of music training in children. For example, in a longitudinal study by Hyde and colleagues (2009) children who underwent a music training showed areas of greater relative voxel size after 15 months of training, with respect to children who did not undergo the music training. These changes were observed in the motor hand area, in the corpus callosum and in a right primary auditory gyrus. It is to note that before the training the children of the music group and the children of the control group did not show any brain difference (Hyde et al., 2009). Corpus callosum was found to be larger in musicians with respect to nonmusicians also by Schlaug and colleagues (1995), and the authors concluded that this might be a product of learning and practicing coordination between left and right hand (since the corpus callosum connects the left and the right hemisphere in the brain). Interestingly, this difference was found only in musicians who started the training before the age of seven (maybe because of childhood periods more sensible to plastic changes) and only in males (hormones could play a role) (Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995). From a functional point of view, musicians, in comparison to nonmusicians, show increases in neuronal auditory responses when hearing unexpected tones in a melody (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004a). Moreover, the cortical representation of the fingers is enhanced in musicians: for instance, the representation of left hand fingers in violinists is expanded in comparison to their right hand and to controls (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995).

All these studies enable to realize the importance of the training in shaping our brain, and, of course, our behavior too. From a behavioral point of view there are studies that compared experts and non-experts in expertise-related tasks, but also in other general cognitive tasks. Concerning expertise-related tasks, some studies observed how expert chess players can remember complex patterns of chess positions after only few seconds of viewing, something that novice players cannot do easily (e.g., deGrott, 2014; Chase & Simon, 1973). Some other studies focused instead on athletes, such as skilled soccer players, observing how these athletes have better anticipation and decision-making skills (Williams & Davids, 1995). Namely, they can anticipate better the actions of their opponents, they have better visual search abilities and they can predict with more precision what is likely to happen in a given circumstance. These tasks, anyway, recreate the situation of a football match (Williams, 2000).

These behavioral characteristics of experts do not surprise: if somebody is an expert, for sure s/he will perform better in their domain of expertise with respect to non-experts. What it is more interesting, in contrast, at a behavioral level, is when (and if) there is a generalization (also called “far transfer”) of skills to a domain not directly trained by the expert. For example, a meta-analysis by Voss and colleagues (2010) found a medium effect size in processing speed and attention tasks with athletes performing better than non-athletes (Voss, Kramer, Basak, Prakash., & Roberts, 2010). Another recent study on tennis players, observed how these class of athletes performed better than non-athletes in a Stroop task (Pacesova, Smela, Kracek, Kukurova, & Plekova, 2018). This task is used to tap information-processing speed, selective attention and inhibitory control. This is not strictly related to their training (mostly physical), therefore it could be a case of far transfer, if we are assuming that it is the training the cause this enhancement. In any case, there are several factors implied in doing sports that could be responsible of these results. As claimed by the authors, these factors could be learning to cope with stress (and therefore performing better in difficult tasks that enhance stress) or spending time in an activity that engages executive functions (e.g., decision-making).

Moving away from athletes, we can find other behavioral studies on experts, for example, in the language domain. In a study by Christoffels, Groot, and Kroll (2006), language interpreters were compared with bilingual university students and with language teachers, and they found that the former performed better in several working memory tasks (e.g., reading span, word span) with respect to the other two groups. However, one could claim that interpreters possessed already better working memory skills before becoming interpreters, and that it is not a mere consequence of the training they underwent. Anyway, results of other studies about trainings and brain plasticity that produce functional and structural changes even after few sessions (Draganski et al., 2004, Pascual-Leone et al., 1995, Pascual-Leone, Amedi, Fregni, & Merabet, 2000, see Slagter, Davidson, & Lutz, 2011 for a review) suggest the peculiar and intense training that interpreters do might produce changes that reflect also their working memory capacity. A major problem of this kind of studies is that it is difficult to tell whether the enhanced skills are a consequence of the specific training or if the individuals who underwent the training were already “better” (i.e., more efficient cognitive skills). This is one of the biggest problems of the studies that compare experts and non-experts: they are all quasi experiments, which means that they do not allow to infer if there is a causality relationship between two variables (e.g., the training and the enhanced skill).

To conclude, even if quasi experiments have several limits (which will be further discussed at the end of chapter three), it is quite clear that expertise is associated to structural and functional differences in the brain, and to some extent, also to cognitive differences (note that this association does not mean that there is causality). Concerning cognition, though, the evidence is weaker when considering domains not strictly connected to the area of expertise; there are studies that did not find any difference between experts and non-experts at the cognitive level (or they found it partially), especially when considering expert athletes (e.g., Furley & Memmert, 2010; Verburgh, Scherder, Van Lange, & Oosterlaan, 2014; Alves, et al., 2013). The literature is wider and more consistent, instead, when musicians (and nonmusicians) are taken into account. Therefore, the rest of the dissertation will be dedicated to one only class of experts: the musicians.

CHAPTER 2. THE MUSIC TRAINING

At the present point, it is clear that experts differ from non-experts both anatomically and functionally, and also behaviorally. Musicians are a class of experts largely investigated because of their peculiar expertise. In fact, learning to play an instrument requires to integrate several sensory inputs and motor output: the person learns to translate a symbol (i.e., the note), into a fine motor response (bimanual) that is associated with a specific sound. Years after years, the person can produce the right motor response more and more automatically and accurately, being able to play scores of increasing difficulty. This is the technical aspect of the music training, but of course the training involves several other aspects, such as motivation, stress coping, family support, etc. All these aspects together make the music training a unique activity, in which the individual has to be committed, motivated, and supported, in order to succeed and become a musician. Musicians are therefore interesting models, either for looking at the "immediate" effects of the training, such as on the perceptual skills (e.g., discrimination of frequencies) or for looking at its possible association with enhanced nonmusical skills, such as cognitive skills.

In this chapter I will review the main studies on musicians, focusing only on behavioral data experiments, since for the present dissertation only behavioral studies were run. In the first part I will take into account studies about perceptual skills of musicians vs. nonmusicians and in the second part I will focus on studies which investigated cognitive skills of musicians and nonmusicians. I will conclude with a general discussion.

2.1 Perceptual tasks

Several studies investigated the perception of the sound frequency in musicians, or, in other words, the acoustical counterpart of "pitch". There is an association between being a musician, and being able to better discriminate pitches with respect to people who do not play any music instrument (e.g., Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005)

One classic study that investigated discrimination of sounds' frequencies in musicians and nonmusicians is the one by Spiegel and Watson (1984), in which a group of professional musicians and a group of nonmusicians were listening to pairs of tone patterns. In each trial, participants listened to the two patterns and had to judge whether they were identical or different. The authors estimated a threshold of frequency discrimination. The median threshold of musicians was about one third of the size of the nonmusicians' threshold, meaning that they could perceive the difference of the two patterns with smaller frequency differences with respect to nonmusicians (Spiegel & Watson, 1984). In another study professional musicians and nonmusicians were compared again on a frequency discrimination task, in which there were three consecutive pure tones (i.e., the standard, played twice, and comparison), and again the participant had to judge if they were identical or different. Results indicated that musicians performed better than nonmusicians. (Kishon-Rabin, Amir, Vexler, & Zaltz, 2001). Another dimension of music which was investigated in musicians and nonmusicians is rhythm. Musicians can, for example, reproduce rhythmic patterns more accurately than nonmusicians (e.g., on a drum, Drake, 1993, or by tapping, Franěk, Radil, Beck, & Pöppel, 1991).

Pitch and rhythm are music dimensions; therefore, it is not surprising that musicians perform better than nonmusicians. As mentioned in the first chapter, it is interesting to investigate whether musicians perform better also with aspects that they do not train directly. For example, in a study by Parbery-Clark and colleagues (2009), a group of musicians performed better than a group of nonmusicians in a speech-in-noise recognition task. In this task, participants had to repeat simple sentences that were played with background noise. Their superiority could be due to a better ability to segment the speech, and to a better use of the cues for understanding the words. It is not clear, again, if the music training is directly responsible of these enhanced skills (Parbery-Clark, Skoe, Lam & Kraus, 2009). These results were replicated also in a following study on children, and moreover, children who underwent a music training were performing better also on measures of auditory attention and auditory working memory (Strait, Parbery-Clark, Hittner, & Kraus, 2012). Note that the superiority of musicians over nonmusicians in the speech-in-noise recognition is not always supported by other studies: for example, Ruggles, Freyman and Oxenham (2014) did not observe this advantage.

Other interesting studies observed how musicians can better discriminate the prosody of speech, even in a foreign language (Marques, Moreno, Castro, & Besson, 2007), and also how they can better recognize the emotions it conveys. For example, in a study by Lima and Castro (2011), musicians and nonmusicians listened to sentences with a neutral semantic content, that were presented with different prosodies. Participants had to judge which emotion was conveyed by the prosody, on a four forced-choice scale. Musicians showed higher accuracy in the recognition of emotion than nonmusicians. These studies show an example of how musicians perform better in tasks that are not strictly related to the music domain, but that are part of a more general auditory perception domain. It seems that this better performance is driven by the music training (see Kraus & Chandrasekaran, 2010 for a review). Now we will make a further step towards a domain that is even farther than the perceptual one from the music training: cognition.

2.2 Cognition, music training, and music aptitude

Cognitive skills are involved in any kind of human activity, and playing a music instrument surely engages several cognitive abilities, such as memory, and executive function (e.g., attention, planning, controlling). What it is interesting, though, is that musicians seem to possess enhanced cognitive abilities with respect to nonmusicians, and not only when performing tasks in the music domain (e.g., tone recall, Williamson, Baddeley, & Hitch, 2010; working memory for tonal and atonal sequences, Schulze, Dowling, & Tillmann, 2012). Is the music training responsible of these enhancements? Or were musicians already skilled before starting the training? This is the recurrent question that quasi experiments cannot answer. However, since quasi experiments are the most numerous studies on cognition in musicians, I will now review some of them. In addition, I will include also longitudinal studies (that unfortunately are a few) that instead tried to investigate the causality relationship between music training and enhanced cognitive skills. I will describe literature separately for cognitive functions investigated: language, visuospatial skills, executive functions, and general intelligence. Music aptitude will be also considered, because a relationship between music perception skills and cognition was observed in several studies. An entire chapter on memory will then follow.

2.2.1 Music training and language skills

The relation between music and language has been widely investigated, in fact, both are universal forms of communication and they are both composed by sounds' sequences. There are some authors that claim that humans process (at a cerebral level) music and language differently (e.g., Peretz et al., 1994; Peretz & Coltheart, 2003), whereas other authors claim that the two domains overlap in terms of networks of processing (e.g., Patel, 2003; Schön et al., 2010). Some behavioral studies support the fact that language and music use common processes, because, for example, musicians show often better language skills (and therefore, training in music would enhance also their verbal skills). Note that this could be also explained by a "far-transfer" of the skills trained by musicians to the verbal domain, even if there is not overlap at a brain level. The debate is still open, but in any case, it is interesting to note that musicians do perform better in some language tasks (e.g. on vocabulary tasks, Piro & Ortiz, 2009). For example, a study by Moreno and colleagues (2009) investigated reading skills in a longitudinal study on children. Two training programs (painting and music) were assigned to two different groups of children, for a period of six months. A word reading test was administered before and after the training. The music group was the only one that improved in the word-reading test. Concerning reading abilities, a positive association between phonological awareness (which predicts reading ability) and music training was found in children after only four months of music lessons (Gromko, 2005).

Another dimension of language which was investigated is reading comprehension. Corrigall and Trainor (2011) found a positive association between the length of a music training and reading comprehension skills, meaning that children with longer music training had better comprehension skills than children with shorter music training, even when age was held constant. Other quasi experiments focused on the second language skills. Some studies, for example, observed how adult musicians are better than nonmusicians in discriminating the phonemes of a foreign language (e.g., Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Martinez-Montes et al., 2013). To conclude, literature suggests that music training is positively associated to enhanced language skills, and some studies in particular suggest that this enhancement might be due to the training itself (Moreno et al., 2009; Corrigall & Trainor, 2011).

2.2.2 Music training and visuospatial skills

Playing a music instrument involves also visuospatial skills, because the notes on the instruments are spatially coded, and the person learns sequences of movements on different positions on the keys. It seems reasonable that musicians will train their visuospatial skills needed for their instrument, but what about general visuospatial skills (i.e., not linked to the instrument played)? One test often used to tap visuospatial abilities is the mental rotation test. In a study by Sluming and colleagues (2007), musicians and nonmusicians had to complete a mental rotation task, in which pairs of 3D drawings were presented, with one drawing of the pair which was rotated with respect to the other. The participants had to judge whether the pair was composed by two identical drawings, or if they were different. Musicians performed better than nonmusicians in this task, and neuroimaging data also supported these superior visuospatial abilities in musicians (Sluming et al., 2007). In another study by Stoezs and colleagues (2007), musicians were performing better than nonmusicians in a visual-search task, an embedded figure test, in which they had to detect a target figure in a complex line drawing (Stoezs, Jakobson, Kilgour, Lewycky, 2007). Musicians seem to perform better than nonmusicians also in a line orientation task (Sluming et al., 2002).

Again, the above studies are quasi-experiments, so it is not really clear whether this enhancement in visuospatial abilities in musicians could be an effect of the music training. Nonetheless, there are some longitudinal studies on children that suggest it might be the music training that drives this enhancement in visuospatial abilities (Rauscher & Zupan, 2000; Rauscher, 2002)

2.2.3 Music training and executive functions

One can argue that the advantage in several cognitive task, such as visuospatial and language task, is mediated by executive functions, that could be enhanced in musicians (and therefore, there should be no specific advantage but a more general one). Executive functions in fact, are involved while playing a music instrument, because they are responsible of controlling, inhibiting, planning, applying strategies, directing attention, and so on. If these functions are enhanced in musicians, they could allow to perform better in any cognitive task. Some authors believe that executive functions might mediate the advantage seen in musicians in cognitive abilities (Schellenberg & Peretz, 2008).

For example, intelligence and executive functions were assessed in children from 9 to 12 years old, taking into account children who were undergoing a music training and children who were not. A positive correlation between years of music lessons and intelligence was found, but this relationship was mediated by executive functions, in particular by attention and by inhibition (Degé, Kubicek, & Schwarzer, 2011). Another study by Bialystock and DePape (2009), musicians were outperforming nonmusicians in a Stroop task, which presented a conflict between pitch and word, meaning they had better control and inhibition; they also performed better on a task that involved a conflict between a target cue and its position, revealing again better control also on a nonverbal spatial task.

Better performance of musicians over nonmusicians was found in visual attention tests, specifically, in selective, divided, and sustained attention (Rodrigues, Loureiro, & Caramelli, 2013). If attention is enhanced in musicians, it could be one of the reasons they can perform better in cognitive tasks. Nevertheless, some studies did not find any superiority and/or mediation effect of executive functions in musicians on other cognitive tasks (Schellenberg, 2011; Costa-Giomi, 1999). Even though the idea that musicians have better cognitive abilities thank to their enhanced executive functions is interesting, there are not enough data to confirm this hypothesis.

2.2.4 Music training and general intelligence

Some studies observed that musicians have a higher general intelligence with respect to nonmusicians, and this could explain why they perform better in several tasks. One of the most famous examples is the longitudinal study by Schellenberg (2004), in which he assigned randomly 144 children to music classes, drama classes and no classes (control group). He administered the WISC (Wechsler Intelligent Scale for Children, Wechsler, 2003) to have a measure of intelligence, before the training and after one year of training. Only the music group significantly improved in the intelligence test. This could suggest that music training enhances general cognitive functions, but there are not many evidences in support (i.e., lack of true experiments)

In general, a correlation between music training and intelligence was found in other studies, for example, Schellenberg later in 2006 found a positive association between duration of music training and intelligence, and there was not any specific association between music training and cognitive abilities when intelligence was held constant, suggesting that any specific advantage depend on a higher intelligence (Schellenberg, 2006). Other studies observed that musicians have higher intelligence than nonmusicians (i.e., higher scores on intelligence standardized test, e.g., Schellenberg, 2011; Schellenberg & Mankarious, 2012; Gibson, Folley, & Park, 2009) and only a few studies did not find any difference between groups (Brandler & Rammsayer, 2003), or found an advantage in specific skills when holding constant the IQ score (e.g., reading comprehension: Corrigal & Trainor, 2011; memory: Jakobson, Lewycky, Kilgour, & Stoesz, 2008). One of the problems is that often quasi experiments have no (or limited) control measure of general intelligence therefore it is not clear if intelligence has a role in the advantage of musicians in cognitive abilities or not.

2.2.5 Music aptitude and cognitive skills

Music aptitude can be defined as a natural ability to perceive and discriminate music. It is an interesting aspect to consider because it can be found in the population along a continuum; musicians can have higher or lower music aptitude, and nonmusicians too, could possess a good music aptitude, though they never underwent a music training. What is it like to have a high music aptitude? For example, and individual with a good music aptitude will be able to discriminate subtle changes in melodies, to detect out of pitch tones, to perceive small rhythmic changes, and other fine discriminations of music features. Possibly, an individual who possesses a good music aptitude will learn with less effort to play an instrument, but there is not yet experimental evidence of this assumption. In any case, having a high natural music aptitude (without having received any training) is not essential for learning to play an instrument (e.g., and individual can always train his/her music perception skills with practice), but being a musicians it is usually associated to higher music aptitude (e.g., Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010), though it is not clear if it is a consequence of the training or not. Probably it is, giving the fact that perceptual skills can be trained and improved with a training(Micheyl, Delhommeau, Perrot, & Oxenham, 2006). It is interesting to investigate music aptitude because it overcomes the usual separation between musicians and nonmusicians (with all the limits of quasi-experiments), and it allows to look directly at the relationship between musical skills (in this case not practical but perceptual) and cognitive abilities.

Music aptitude can be investigated with tests that are usually based on a recognition paradigm. In these tests the individual listens, for instance, to a melody, and then after few seconds s/he listens to another melody and has to judge whether the second melody is the same or different from the first. The stimuli could be other musical feature, such as rhythmic patterns, and chords. The first music aptitude test is that by Seashore (1915). This test, which is still used, is composed by six different subtest that tap various aspect of music perception, such as pitch, loudness, timbre, rhythm, memory for tones and meter (i.e., speed) (Seashore, Lewis, & Saetveit, 1960). One of the most recent tests is the one by Law and Zentner (2012), that it is called PROMS (i.e., Profile of music perception skills), and it will be described in detail in STUDY 2.

Several studies investigated music aptitude in relationship with nonmusical abilities. Concerning language skills, for example, individuals with high music aptitude perform better than individuals with lower music aptitude in phonological awareness tasks. Phonological awareness can be defined as the ability to perceive and segment the sounds that compose a word (Blachman, 2000), and these tasks require, for example, to repeat a word excluding one phoneme (e.g., the first). Children with higher music aptitude perform better in this task than children with lower music aptitude (e.g., Norton et al., 2005; Peynircioğlu, Durgunoğlu, & Úney-Küsefoğlu, 2002). Music aptitude was found to be positively correlated also to reading abilities (Anvari, Trainor, Woodside, & Levy, 2002). A subtest of a music aptitude test that requires to judge how many notes are presented simultaneously was found to be positively correlated with reading ability in children, even when controlling for IQ and age (Barwick, Valentine, West, & Wilding, 1989). Interestingly, some studies observed that deficits in rhythmic aptitude predict problems in phonological awareness and reading in dyslexic children (Goswami, Huss, Mead, Fosker, & Verney, 2013; Overy, Nicolson, Fawcett, & Clarke, 2003). Positive association between music aptitude and language skills were found also in the learning of a second language, in particular in the pronunciation and discrimination skills (Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008; Milovanov, Pietilä, Tervaniemi, & Esquef, 2010; Slevc & Miyake, 2006). Nevertheless, in one of our studies we did not find any relationship between music aptitude and second language abilities, but we found instead a positive association between music training and second language skills (in particular with a dictation task, which relies also on phonological perception) (Talamini, Grassi, Toffalini, Santoni, & Carretti, 2018).

Music aptitude seems to have a positive relationship also with general cognitive abilities. For example, some studies found that children with a greater aptitude for music were also of higher general intelligence (Doxey & Wright, 1990; Norton et al., 2005). In the study by Norton and colleagues, 2005, two groups of children between 5 and 7 years of age were compared on several cognitive measures. One group was composed by children who were beginning music lessons and the other group was composed by children that were not beginning any training. The tests administered were about visuospatial and nonverbal reasoning, verbal, motor, and music perception skills. Magnetic resonance imaging was also collected. The authors found no difference between groups in terms of neural images, motor and music perception skills. However, they found a correlation between music perceptual skills and both nonverbal reasoning and phonological awareness. Authors concluded that music perception skills are connected to visual pattern recognition (that is, the nonverbal reasoning test, the Raven matrices), and to language processing (Norton et al., 2005). Giving the fact that some studies found a positive relationship between music aptitude and IQ, it is not clear whether relationships between music aptitude and specific skills (i.e., language) are only a by-product of a higher IQ. Nevertheless, some studies, as we saw, found a positive association with specific skills even when controlling for IQ (e.g., Barwick et al., 1989). If music aptitude is linked to intelligence, it is not clear why, for example, people with amusia perform poorly on music aptitude but have a normal general cognitive abilities (e.g., Peretz et al., 2002).

2.2.6 Summary and conclusions

Music training is associated to enhanced auditory perception skills, even with nonmusical material. It is possible that the training improves the ability of perceiving and discriminating sounds, since a “good ear” it is necessary to play an instrument. Interestingly, musicians perform better than nonmusicians also in several cognitive tasks, such as language and spatial skills. Even though it is reasonable to think that the peculiarity of the music training might enhance, for example, attention, visuospatial abilities, and so on, there is yet a little evidence of causality. Moreover, it is not clear if a possible advantage in cognitive skills is mediated by executive functions, or if the advantage can be considered separated for each skill, and/or if simply musicians possess higher general intelligence than nonmusicians.

Another way of looking at the relationship between music skills and cognitive skills is considering music aptitude, which was found also to be positively associated to some cognitive skills, such as language and nonverbal reasoning. Music aptitude allows to understand whether it is the music training which is associated to enhanced cognitive skills, or whether, independently from the training, having a high music aptitude is linked to other enhanced cognitive skills. Even here, it is difficult to tell whether the association found is a consequence of a higher intelligence in general, or if it is specific to some cognitive skills. Regardless of these limits, surely music aptitude is an important aspect that has to be considered when investigating the relationship between music (training, or skills) and cognition.

In this chapter I deliberately omitted memory (which is the main focus of this dissertation), because it will be described in detail in the next chapter, where I will review the main literature.

CHAPTER 3. MEMORY AND MUSIC TRAINING

3.1 Theoretical overview

A premise is necessary when talking about memory: first of all, there is not yet a unique model that describes memory in humans. In the present dissertation I will use a common distinction between long-term memory, short-term memory and working memory (Cowan, 2008).

Long-term memory is the system that allow to store information for few minutes to potentially a whole life-time (Aktinson & Shiffrin, 1968); it can be investigated, for example, with tasks that have a learning phase and a subsequent recall or recognition of the material previously learnt. Short-term memory is the system responsible of retaining information (usually from 5 to 9 items) to for a few seconds up to about half a minute. It is a passive storage, and the person does not have to make any effort (apart from rehearsing) to keep the information in mind (Aktinson & Shiffrin, 1968, Cowan, 2008). It can be assessed with recall tasks, such as forward span tasks (e.g., digit span forward, word span forward, Corsi visuospatial span forward), in which the participant listens or reads a sequence of items and after that s/he has to immediately recall the sequence in the same order of presentation. It can be assessed also with recognition tasks, in which the participant has, for example, to compare two sequences of items and judge whether the second sequence is identical or different from the first one. Working memory is the memory system that enables to maintain a limited amount of information temporarily, manipulating this information at the same time, or performing a secondary task. One example of a task that requires to manipulate the information is the backward digit span, in which the participant has to recall a series of digits presented starting from the last item and proceeding backwardly. An example of a test with a secondary task is the operation span task which requires to see simple arithmetical operations, each one followed by a word, judge whether the solution of the operation is correct or not, and then recall all the words that followed each operation in the right order. Often, working memory and short-term memory are considered the same memory system, but in the present dissertation, the two systems will be differentiated: short-term memory only allows to retain information, and working memory is a more active function that allows to temporarily store process information at the same time (Baddeley & Hitch, 1974; Case, Kurland, & Goldberg, 1982).

Within each memory system, we can also make a distinction basing on the type of stimuli use in the task (e.g., verbal, such as words and numbers; visual, such as figures; spatial, such as spatial positions; musicals, such as tones, melodies, and chords). Moreover, the temporal distinction (i.e., how long we can remember an information) of the various memory systems is not the only possible one; on the contrary, it is important to distinguish also between auditory and visual memory, (and within the auditory domain we can also make a distinction between verbal and musical memory). There is an open debate on the distinction between visual and auditory memory (Fougnie & Marois, 2011), and it is not yet clear whether there are separate "storages" in memory depending on the sensory modality of the input, regardless of the material that has to be remembered (e.g., verbal, spatial). Since the same type of material (e.g., verbal) seems to be processed differently depending on the sensory modality (i.e., visual or auditory, see Crottaz-Herbette, Anagnoson, & Menon, 2004) and giving other results about musicians which found a difference between auditory and visual memory, (Cohen, Evans, Horowitz, & Wolfe, 2011; Talamini, Carretti, & Grassi, 2016) I will describe studies basing also on the sensory modality of the presented stimuli and the type of material to be remembered.

The model of working memory by Baddeley and Hitch (Baddeley, 2000; Baddeley & Hitch, 1974) made a distinction based on the type of stimuli to be remembered. In fact, in his multicomponent model, there is a visuospatial sketchpad, which is responsible of memorizing and manipulating visual and spatial information, and there is a phonological loop, which instead is responsible of remembering and manipulating verbal stimuli. These two systems are controlled by the central executive, which is responsible of controlling and spreading the resources needed by the two sub-systems, and the episodic buffer is another component that connects information between working memory and long-term memory, and between the two subsystems of working memory. The episodic buffer is thought to be also responsible of binding processes, which allow to integrate information coming from different sources with the aim of creating new representations (Baddeley, 2000). For a graphical representation of the model, see figure 5 "The multi-component working memory revision" in."Working memory: Looking back and looking forward" (Baddeley, 2003).

This model was useful for understanding that verbal and visuospatial stimuli are coded differently, but it is likely that the distinction is not so simple. What about musical memory for example? Some authors suggest that there are two separate subsystems: a phonological loop for verbal information and a tonal loop for music information (Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011). In fact, a in study by Schulze and colleagues (2011) which investigated the neuronal correlates of working memory, the authors observed an activation in the same areas in both a tonal and a verbal memory task, but also specific separate activations in musicians when rehearsing tonal and verbal information. Moreover, the structures involved in the tonal task in nonmusicians were all involved also in the verbal task, and this was true only for the nonmusicians (i.e., there was an overlapping of structures involved in the two tasks). Authors concluded that musicians might engage new structures for processing tonal stimuli, due to their training and consequently use two different subsystem for rehearsing tonal stimuli and verbal stimuli (Schulze et al., 2011). Stronger evidence of two separate systems comes from studies with patients with congenital amusia. Amusia is an impairment of music perception related to pitch perception. For example, amusics are unable to detect pitch changes in melodies, to recognize familiar tunes without the lyrics, to detect out of tune and dissonant chords in classical music (Peretz, Cummings, & Dubé, 2007).

Patients with congenital amusia showed a selective disorder for tonal short-term memory and memory for timbre, whereas verbal short-term memory was intact (Tillmann, Lévêque, Fornoni, Albouy, & Caclin, 2016; Tillmann, Schulze, & Foxton, 2009). In another study that required to remember tonal and atonal sequences, again amusia patients showed a deficit, whereas in the verbal digit span they performed normally (Albouy, Schulze, Caclin, & Tillmann, 2013). Individual results coming from a study on stroke patients showed how some of the patients had selective deficits either for tonal short-term memory either for verbal-short term memory, but this dissociation was not due to different hemisphere lesions (e.g., left for verbal deficits, right for musical deficits), suggesting that the networks involved in musical and tonal short-term memory are distinct at some level, but they both underlie on both hemispheres (Hirel et al., 2017). Taken together, these results suggest a separation between verbal memory and musical memory, at least, in short-term memory tasks, meaning that there could be two different temporary storages for these two different materials.

Given these theoretical notions, in the present chapter I will review the literature about music training (by music training I will refer to the long-term music training that a person undergoes to become a musician) and memory skills, separated for memory system; I will start with long-term memory, then moving to short-term memory, and concluding with working memory. Verbal and musical memory will be treated as separate systems, as for, of course, visual and spatial stimuli.

3.2 Long-term memory and music training

Literature about music training and long-term memory is scarce. Nevertheless, there is evidence of a superiority of musicians over nonmusicians in verbal (i.e., words and/or numbers) learning and recall tasks, both in studies with children and with adults (Franklin et al., 2008; Ho, Cheung, & Chan, 2003, Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Cohen et al., 2011; Huang et al., 2010; Roden, Kreutz, & Bongard, 2012; Taylor & Dewhurst, 2017). For example, in the study by Franklin and colleagues

(2008), musicians and nonmusicians were compared on a learning-recall task. Participants listened to a series of 15 words (presented four times) that were read aloud by the experimenter, and at the end of each presentation, they had to recall as many words as possible in any order. Participants had to recall the list also after 30 minutes of delay. The authors had an additional condition with articulatory suppression, a technique used to avoid the mental rehearsal of the words presented: in this condition, participants had to say aloud the word "the" between the words that were presented by the experimenter. Results showed that musicians performed better than nonmusicians (i.e., they remembered a larger number of words) in both immediate and delayed recall, but this advantage disappeared when the two groups were performing the articulatory suppression. The authors concluded that the advantage of musicians might be therefore linked to a better use of rehearsal strategies.

Although there is evidence of a superiority of adult musicians over nonmusicians in verbal long-term memory tasks some studies did not find any advantage (Helmbold, Rammsayer, & Altenmueller, 2005; Brandler & Rammsayer, 2003; Suárez, Elangovan, & Au, 2016). For example, Suárez and colleagues (2016), administered several memory tests to a group of musicians and a group of nonmusicians. Specifically, a nonword recognition task was adminsitered, in which partipants had to learn 24 nonwords that were presented auditorily. After the learning phase, particpants had to judge among a set of nonwords whether they were nonwords previously heard or not. In this test, musicians and nonmusicians performed equally well.

Evidence is even weaker when the stimuli presented in the memory task are visual (e.g., figures); in fact, only one study found that adult musicians performed better than nonmusicians (Jakobson et al., 2008). In this study, participants had to learn 15 simple line drawings that were presented sequentially for five times. After each presentation, participants had to draw as many drawings as possible. After the five presentations, a recognition task for the drawings was administered. Finally, after about 15 minutes from the learning trials, participants had to recall again the drawings previously learnt. Musicians performed better than nonmusicians in both the recall (on the last two trials of the learning part, and on the delayed recall) and recognition tasks (Jakobson et al., 2008).

To the best of my knowledge, this is the only evidence of a superiority of musicians in visual long-term memory; in fact, the remaining studies did not observe any difference between musicians' and nonmusicians' performance, neither in adults nor in children (Ho et al., 2003; Cohen et al., 2011; Roden et al., 2012; Brandler & Rammsayer, 2003; Helmbold et al., 2005; Chan, Ho, & Cheung, 1998). In the study by Cohen and colleagues (2011), musicians and nonmusicians were exposed to several visual objects and abstract art pieces and then complete a recognition task. There was no difference between the two groups, and the advantage of musicians was confined to auditory stimuli (both speech and environmental sounds see Cohen et al., 2011).

There are also studies that investigated memory for musical stimuli, such as familiar and unfamiliar pop songs. For example, always in the study by Cohen and colleagues (2011) authors found that musicians performed better than nonmusicians in remembering familiar and unfamiliar music. Nevertheless, another study involving a melody learning and recognition did not observe any difference between musicians and nonmusicians (Schiavo & Timmers, 2016).

3.3 Short-term memory and music training

Many studies investigated the short-term memory of musicians and nonmusicians, either directly, either indirectly, because it was often used as control measure for other tasks. Results indicates that musicians of all ages perform better than nonmusicians in verbal tasks (such as the span) that require to reproduce sequences of numbers, letters, or words (Anaya, Pisoni, & Kronenberger, 2016; George & Coch, 2011; Hansen et al., 2013; Lee, Lu, & Ko, 2007; Ramachandra, Meighan, & Gradzki, 2012; Roden, Grube, Bongard, & Kreutz, 2014; Suárez et al., 2016; Talamini et al., 2016; Weiss, Biron, Lieder, Granot, & Ahissar, 2014; Tierney, Bergeson-Dana, & Pisoni, 2008).

For example, in the study by Hansen and colleagues (2013), expert musicians, amateur musicians, and nonmusicians were compared on several short-term memory tasks, among which there was also a digit span forward. In this task, expert musicians outperformed nonmusicians, whereas there was no other significant difference among the other groups (i.e., experts vs. amateurs, amateurs vs. nonmusicians). Qualitatively, the performance of amateur musicians was halfway with respect to the other two groups. Moreover, the authors administered a test to tap the music aptitude of all participants. A positive correlation between one of the subtests of the music aptitude test and the score in the digit span forward was found. This subtest requires to judge whether two short rhythmic phrases (that are played sequentially) are the same or not. This subtest positively correlated also with the total score of the music aptitude test, which was composed by another subtest of melody recognition (Hansen et al., 2013).

In the study of Tierney and colleagues (2008), authors manipulated the modality of presentation of the stimuli (i.e., auditory and visual) to see whether a possible advantage in short-term memory of musicians was general or specific to a sensory channel. Four groups of participants participated in the experiment: experienced students of a music school, gymnasts, psychology students, and videogame players. The task used to tap short-term memory was a modified version of the Simon Memory Game, a box with four panels of four different colors connected to a computer. In the visual presentation, the panels were illuminated randomly one at a time, and participants had to reproduce the sequence by pressing the panels. In the auditory condition, participants heard the name of the colors of the panels and they had to reproduce the sequence always by pressing the panels. Finally, there was also an audiovisual condition in which the two previous conditions were combined. Results showed that musicians performed better than the other two groups in the auditory condition, but no difference among groups was found in the other two conditions. Authors concluded that the advantage of musicians over nonmusicians is likely to be selective for auditory stimuli, and not driven by other variables such as, for example, IQ. Even if the results of several studies support an advantage of musicians over nonmusicians in verbal short-term memory tasks, two studies with adult musicians and nonmusicians did not observe any difference between groups (Boebinger et al., 2015; Okhrei, Kutsenko, & Makarchuk, 2017). For example, Okhrei and colleagues presented three short-term memory tasks, in which the subjects saw sequences of consonant or of digits (depending on the condition) and then had to judge whether a following item was present in the sequence previously seen or not. The authors did not observe any difference between groups in any of the two conditions.

Concerning visual and spatial stimuli, differently from long-term memory here some studies found that adult and older musicians outperformed nonmusicians in tasks (such as the span) that demand to reproduce visual and spatial sequences (Amer, Kalender, Hasher, Trehub, & Wong, 2013; Bidelman, Hutka, & Moreno, 2013; George & Coch, 2011; Suárez et al., 2016; Yang, Lu, Gong, & Yao, 2016). For example, Suàrez and colleagues (2016) found that musicians performed better than nonmusicians in a task that required to look at a matrix, in which some lines were turning red, one at a time, and reproduce the sequence of the lines turned red in the same order of presentation (Suàrez et al., 2016). Nevertheless, the results of several other studies did not support this advantage (e.g., Bialystok & DePape, 2009; Hansen et al., 2013; Okhrei et al., 2017; Rodrigues, Loureiro, & Caramelli, 2014; Tierney et al., 2008). For example, in the same study of Hansen and colleagues (2013) previously mentioned, in the measure of spatial span forward, which requires to reproduce a spatial sequence showed by the experimenter by tapping some cubes on a board (such as the Corsi test), no difference between groups was found.

Finally, with musical stimuli such as tones, chords, and melodies, with no surprise adult musicians performed better than nonmusicians (Bidelman et al., 2013; Monahan, Kendall, & Carterette, 1987;

Pallesen et al., 2010; K. Schulze, Mueller, & Koelsch, 2011; Williamson, Baddeley, & Hitch, 2010; Schulze, Dowling, & Tillmann, 2012).

3.4 Working memory and music training

We saw that in short-term memory there is often an advantage (especially with verbal stimuli) of musicians over nonmusicians, but what about more complex tasks? A lot of studies were interested in understanding whether the working memory of musicians is more efficient of the one of nonmusicians. They observed that musicians performed better than nonmusicians in tasks that require to store and manipulate verbal information or recall information while completing a secondary task (Bergman Nutley, Darki, & Klingberg, 2014; Clayton et al., 2016; Franklin et al., 2008; George & Coch, 2011; Hanna-Pladdy & Gajewski, 2012; Lee et al., 2007; Mandikal Vasuki, Sharma, Demuth, & Arciuli, 2016; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011; Ramachandra et al., 2012; Roden et al., 2014; Suárez et al., 2016; Talamini et al., 2016; Zuk, Benjamin, Kenyon, & Gaab, 2014).

For example, Clayton and colleagues, in 2016, administered several tasks to tap executive functions and general intelligence in musicians and nonmusicians. Among these tasks, they administered a digit span task backwards, and found that musicians performed better that nonmusicians, whereas in the other executive functions tasks (e.g., inhibition control, cognitive flexibility) no difference between group was found (Clayton et al., 2016). In a quasi-experimental longitudinal study by Roden and colleagues (2014) authors administered several measures of working memory to two groups of children of 7-8 years of age: one was undergoing a music training, and one was receiving extended education in natural sciences. Participants were not randomly assigned to the two groups, but it was the children’s families that decided which program to choose. The experimenters administered the tests three times over a period of 18 months. Children of music group showed greater improvements over time (even when controlling for age, IQ, and socioeconomical status) in two tests that tapped working memory.

If, again, several studies found an advantage, others could not confirm these results, or the advantage emerged only in children but not in adults (Boebinger et al., 2015; Hansen et al., 2013; Lee et al., 2007; Katrin Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011;). For example, in the study by Hansen and colleagues (2013) previously described, there was no difference between groups on the spatial span backward test. In the study by Lee and colleagues (2007), children who underwent a music training performed better in the digit span backward test and in the operation span test than children who did not undergo any training, even when matched on the Raven Standard Progressive Matrices Test, which is considered to be a measure of fluid intelligence, and on their socioeconomical status. Nevertheless, when they administered the same tests to adult participants divided in musicians and nonmusicians, no difference between groups emerged in these two tests (Lee et al., 2007).

When the stimuli presented in the task were visual and spatial, some studies found that musicians outperformed nonmusicians, but these results emerged only in children (Bergman Nutley et al., 2014; Lee et al., 2007;). Instead, with adults or elderly participants there was no difference between musicians and nonmusicians (Bialystok & DePape, 2009; Hanna-Pladdy & Gajewski, 2012; Hansen et al., 2013; Lee et al., 2007; Parbery-Clark et al., 2011). For example, again in the study by Lee and colleagues (2007), children who were undergoing a music training had an advantage over children who did not in a spatial span, which required to remember which object (among five possibilities) was presented in which position. Differently from children, adult musicians did not perform better than nonmusicians in this task (Lee et al., 2007).

Along the same line of long-term memory and short-term memory, when the working memory tasks presented musical stimuli, adult musicians performed better than nonmusicians (Pallesen et al., 2010; Schulze et al., 2012). For example, in the study by Pallesen and colleagues (2010), classical musicians and nonmusicians were performing an n-back task in which they had to memorize the octave some chords. Specifically, participants had to judge if the octave of the chord presented matched that of the previous trial (in the 1-Back task), or if it matched the chord presented two trials back (in the 2-back task). Musicians outperformed nonmusicians in the 1-back task, which requires less cognitive load, but no statistical difference between groups was found in the 2-back task, even if from a qualitatively point of view the performance of musicians was slightly better.

3.5 Limits of the literature

Most of these studies suggest that musicians do have better memory skills that nonmusicians, but we have to be cautioned when drawing conclusions. First of all, most of these studies are quasi experiments; as already mentioned, this does not allow to understand whether the music training is a cause of these enhancement or a consequence (in the sense that people already “gifted” are more likely to undergo a music training). Nevertheless, most of studies (but not all) tried to use some control measure for general cognitive abilities or IQ. The problem is that there is a large variability in the choice of these control tasks, maybe due to the fact that we have many different cognitive models from a theoretical point of view. It is true, though, that several studies did found an advantage of musicians over nonmusicians even when controlling for these variables. This can suggest that a difference in intelligence, for example, is not responsible of this advantage, but it does not exclude that there are other variables responsible of these results. For example, musicians might be generally more motivated people, who are more likely to commit in doing activities and overcome difficulties, and this can be a useful characteristic also when completing tests in a laboratory. The problem is that this kind of variables are not easy to test. Therefore, it will be always difficult to exclude an a-priori hypothesis (i.e., the difference between musicians and nonmusicians were already there before starting the music training) for a causality hypothesis (i.e., music training enhances cognitive abilities). Longitudinal studies could in part overcome these limits, by, for example, assigning randomly the participants to different training groups. Unfortunately, these kinds of study are not common, probably because of all the difficulties (either practical, that financial) that are inherent in longitudinal studies, and in studies on children.

Another limit of the literature is that musicians are not always defined in the same way. In some studies, musicians were professional orchestra players, whereas in others they were students of a music conservatory. Sometimes the minimum of years of music training is 10 years, whereas in other cases the minimum is lower. There is not a “standard” musician, and this increases the variability across studies and could be also responsible of the different results. Moreover, music aptitude is rarely considered. There are several tests that can be used to investigate music aptitude, and they could be helpful for several reasons: first of all, to have an objective measure to distinguish musicians from nonmusicians. Secondly, to have a continuum of music skills among all participants. Music aptitude can be found in everybody, and it can vary also among musicians. It is true that it is merely perceptual (no playing is involved in these tests), but it could help anyway in understanding if the relationship between music training and cognitive skills is specific to the training itself or if it is linked to a good music aptitude.

Another problem, that does not concern only this kind of researches but also cognitive science in general, is that often cognitive abilities are not well defined. The consequence is that different researchers use different tests, for example, to tap the same ability. All of these limits make it tricky to investigate the relationship between music training and cognitive abilities. Moreover, most of the studies investigated only verbal memory or visual memory, excluding other possibilities, such as auditory memory for nonverbal stimuli, such as environmental sounds (Cohen et al., 2011). In fact, most of verbal tasks were presented auditorily, so that it is difficult to disentangle a general verbal memory advantage from a general auditory memory advantage (not specific for words). The only attempt in administering the same task in different sensory modalities was the study of Tierney and colleagues (2008). Nevertheless, a problem of this study was that the task was not strictly verbal, because each label was associated to a spatial position (for the response), so even the auditory modality could be coded also spatially.

Finally, often the numerosity of the sample is not high, especially in the case of quasi experiments, which do not have randomized samples (i.e., in this case a small number could even more problematic because of possible confounding variables) and this can lead to a low statistical power (Suresh & Chandrashekara, 2012). Nonetheless, these studies enabled to consider music training as an important tool for the development of our brain and cognitive skills, even if there is not yet certainty about the causal relationship, researchers start to acknowledge the power of music and its eventual benefits on us.

CHAPTER 4. THE RESEARCH

The project

My research project investigated the differences between musicians and nonmusicians in memory skills, in particular short-term and working memory. As literature suggests, music training is related to improved memory skills, but there are some studies that did not find any difference between musicians and nonmusicians, and there is a large variability in terms of studies' design.

The first study (STUDY 1) that will be described is a meta-analysis. The meta-analysis was conducted after study 2, but here it will be described before to give a quantitative overview of the literature. The aim of the meta-analysis was to investigate deeply literature, because the effects observed when comparing musicians and nonmusicians in memory skills seemed small and dependent on many variables. For this reason, the meta-analysis (the first on this topic) focused on the studies that investigated the relationship between music training and memory, separately for long-term, short-term, and working memory. Moreover, it took into account the type of stimuli presented in the memory tasks (e.g., verbal, visuospatial). The final goal was to understand if musicians do perform better than nonmusicians in memory tasks, how large the difference between groups is, whether the difference changes as a function of the memory system and of the type of stimuli that are presented in the task.

STUDY 2. The second study is an experimental research that aimed to understand whether the advantage found in verbal short-term and verbal working memory was somehow dependent on the modality in which the task was delivered (i.e., stimuli presented auditorily or stimuli presented visually). In fact, most of studies that tested verbal memory, used tasks that were presented auditorily. Since we know that musicians have better auditory perception skills, this might have helped them in performing better than nonmusicians in these tasks.

In a following study, (STUDY 3), the aim was to take into account individual differences among musicians. In particular, the goal was to understand whether the type of music training (classic vs self-taught) could influence the advantage of musicians over nonmusicians in verbal working memory skills. Specifically, the two groups of musicians differed by one characteristic: the ability of reading a music notation. In this way, the idea was to separate the classic training in which the person learns to translate a symbol (i.e., the note) to a movement, spatial position and sound, from a general ability of playing an instrument without reading. Any possible difference between groups could help understanding which parts of the music training are linked to better memory skills.

STUDY 4 aimed to investigate the difference between visual and auditory short-term memory, using nonverbal and nonmusical stimuli, in order to understand whether the advantage of musicians is linked only to auditory short-term memory or not (as partly suggested by the meta-analysis and the state of the art). No previous study investigated this. Usually the comparison between modalities (i.e., visual and auditory) is made between two different domains (verbal or musical for auditory memory, nonverbal for visual memory). This experimental study was the first to compare visual and auditory nonverbal and nonmusical short-term memory in musicians and nonmusicians. Findings could help understanding why musicians have a better short-term memory than nonmusicians and whether this advantage is only related to auditory memory.

Study 1. Musicians have better Memory than Nonmusicians: A Meta-Analysis

Authors: Francesca Talamini*, Gianmarco Altoé**, Barbara Carretti*, Massimo Grassi*
*Department of General Psychology, University of Padua, Italy.
** Department of Developmental and Social Psychology, University of Padua, Italy

Published on PLOS ONE, October 2017. doi: https://doi.org/10.1371/journal.pone.0186773

Aims

The present meta-analysis had the aim of exploring published studies that were conducted on adult musicians and nonmusicians, in order to clarify whether there is a difference in memory performance and in order to understand how large this difference is. In particular, we wanted to explore and quantify literature separately for long-term, short-term, and working memory. In addition, within each memory system, we wanted to consider the different categories of stimuli presented. Specifically, we wanted to compare memory for verbal stimuli, for visual stimuli, for spatial stimuli, and for tonal stimuli. This distinction was important because it allowed to understand whether the hypothetical advantage of musicians was domain-specific (e.g., in memory tasks with tonal stimuli,), or if it generalized to other domains (e.g., in memory tasks with verbal stimuli, and/or with visuospatial stimuli). Expertise is known to improve domain specific abilities (e.g., Ericsson & Charness, 1994), but if it generalizes to other domains, and to which domains, is not yet clear (see Chapter 1). Moreover, since in literature results are mixed, and since there was not yet a quantitative review, we felt the need of looking deeply inside these studies in order to be able planning future studies at best.

Concerning the age, we decided to focus on young adults because most of studies had participants of this age, whereas studies eligible for the present meta-analysis which involved children and older adults were not many at the time of the collection of the data (i.e., children: N = 7; older adults: N = 2). Moreover, the problem of including children concerns the years of music training, that usually are not many, and make it difficult to compare children to expert adult musicians. On the other hand, as far as older adults is concerned, the performance of older adults varies considerably in comparison to the performance of young adults, therefore the interpretation of the findings would have been more difficult (e.g., Grassi & Borella, 2013).

Method

Study selection

The search of literature was made using the AIRE portal, which is a service provided by the University of Padua. This search tool allows to look for studies across multiple databases: Education Source; PEP (WEB) Psychoanalytic Electronic Publishing; Psychology and Behavioral Science (EBSCO); PsycINFO (Ovid); PubMed; ScienceDirectAllBooks Content (Elsevier API); SCOPUS (Elsevier API); SocINDEX with Full Text (EBSCO); Web of Science. Google scholar was also used in a following step. The key words used to search for studies were: memory, musicians, nonmusicians. Other papers were also found by checking the references of the identified papers. We did not include studies from the grey literature (i.e., unpublished studies), because these studies did not go through a peer review process. In addition, any search into the grey literature will be hardly conclusive, because it is impossible to identify all the relevant unpublished studies available. We are aware that this choice could expose our findings to a publication bias, that is the tendency of publishing only studies that report statistically significant results, and not those with null results); nevertheless, in the present meta-analysis we used some techniques to try to control for a possible publication bias. In the present meta-analysis, the latest search for studies was carried out on February the 15th, 2017.

Inclusion criteria

The criteria we used for including studies were the following: (1) studies with adult participants; (2) studies that included at least two groups of participants: expert musicians (i.e., who had attended music conservatories or music schools), and nonmusicians (i.e., participants who had little to none experience of playing a music instrument); (3) studies that administered a memory task to both groups, that could be classified either as a long-term memory task, a short-term memory task, or a working memory task (see the following paragraph for details of the categorization); (4) studies in which the stimuli used could be classified as verbal, visual, spatial, or tonal; (5) studies published in English.

We selected a pool of studies that matched the criteria just described, and then some exclusions were made because of missing data (data not provided by the authors after we contacted them), and/or because the tasks that were administered were not comparable to those of the other selected studies. The PRISMA flow diagram (Moher et al., 2009) represents all the steps of the literature search (see Figure 1). The quality of the studies was assessed by two independent raters, in particular whether the characteristics of the participants and the tasks administered were adequately described. In case of disagreement, the raters were consulting and discussing the original article in order to solve any issue.

This assessment allowed to screen the studies binarily (i.e., pass or fail) for the inclusion in the analysis.

PRISMA 2009 Flow Diagram

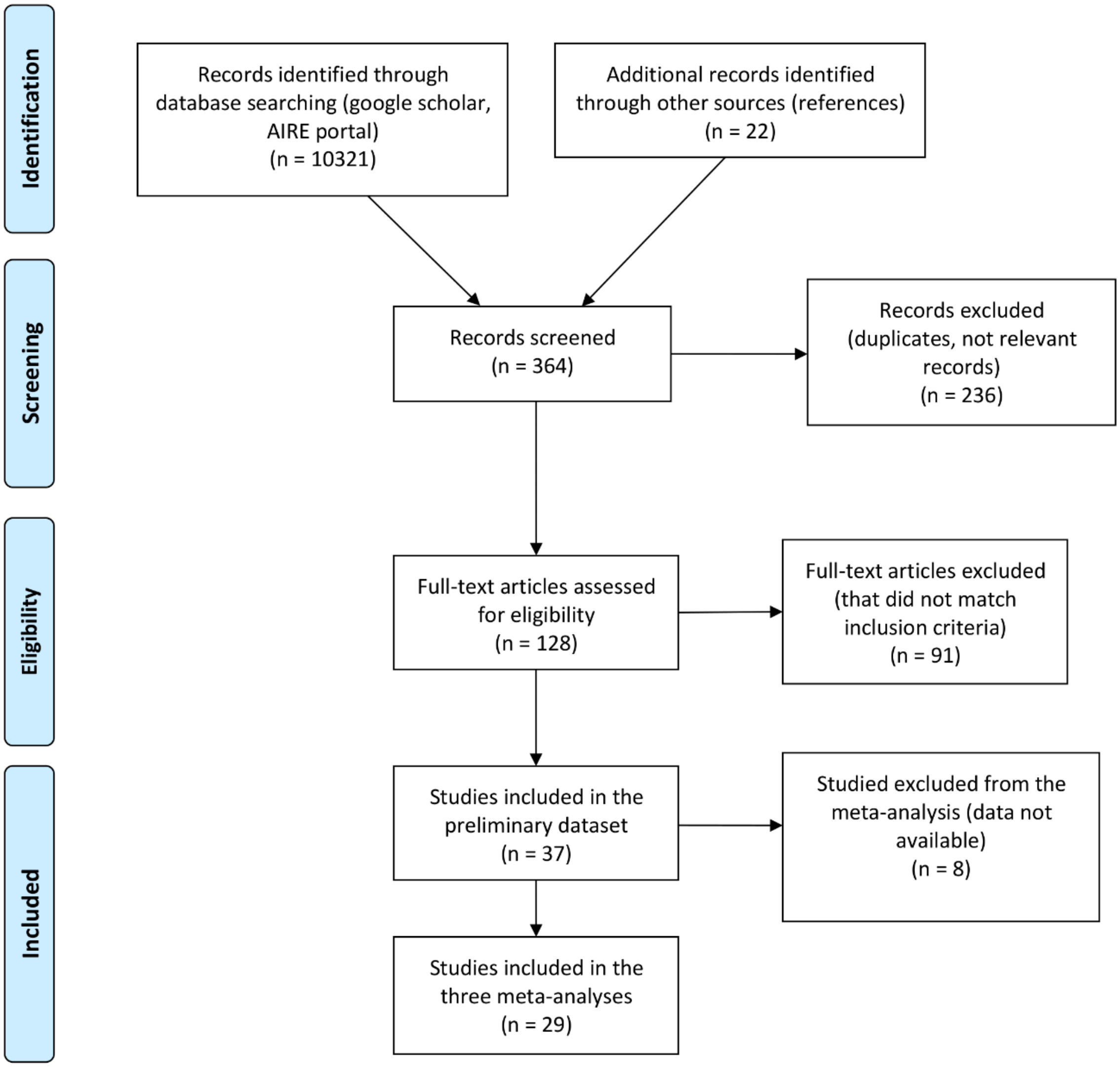

Figure 1. The Prisma Flow Diagram represents the steps of the literature search and screening.

Categorization of memory tasks

We divided the studies depending on the memory system tapped in the experiments, distinguishing between long-term, short-term, and working memory. Thus, studies were classified following this distinction (see Chapter 3 for the description of the memory systems).

Memory tasks were classified as follows. Long-term memory tasks were tasks that required a delayed recall or recognition of information previously learnt in a dedicate phase. Short-term memory tasks included tasks such as the forward span tasks for both verbal stimuli (words and numbers), and visual and spatial stimuli (figures and spatial positions), and recognition tasks when using musical stimuli (tones, melodies, chords). Working memory tasks were tasks that demanded to either perform a secondary task simultaneously to the primary recall task, or to manipulate the information to be remembered (e.g., backward span tasks). Table 1 shows the complete list of tasks classified according to the three memory systems.

Table 1. *List of tasks with the respective memory system tapped*

TASK	MEMORY SYSTEM
Berliner Intelligenzstruktur Test - Recognition of two-digit numbers	LONG-TERM MEMORY
Berliner Intelligenzstruktur Test - Recognition of buildings on a city map	LONG-TERM MEMORY
Berliner Intelligenzstruktur Test - Recognition of previously memorized nouns	LONG-TERM MEMORY
Learning-recall of words lists	LONG-TERM MEMORY
Benton Visual Retention test	LONG-TERM MEMORY
Rey–Osterrieth complex figure test - delayed recall	LONG-TERM MEMORY
The Rey Auditory Verbal Learning Test - delayed recall	LONG-TERM MEMORY
Recognition of previosuly memorized words	LONG-TERM MEMORY
Recognition of previosuly memorized melodies	LONG-TERM MEMORY
California Verbal Learning Test	LONG-TERM MEMORY
Rey Visual Design Learning Test	LONG-TERM MEMORY
Figure recognition	LONG-TERM MEMORY
Non-words recognition	LONG-TERM MEMORY
Wechsler Memory Scale - Visual reproduction II	LONG-TERM MEMORY
Digit span forward	SHORT-TERM MEMORY
Test of Memory and Learning - Digits forward	SHORT-TERM MEMORY
Test of Memory and Learning - Letters forward	SHORT-TERM MEMORY
Test of Memory and Learning - Abstract visual memory	SHORT-TERM MEMORY
Test of Memory and Learning - Memory for location	SHORT-TERM MEMORY
Spatial span forward	SHORT-TERM MEMORY
Nonword span	SHORT-TERM MEMORY
One-back task	SHORT-TERM MEMORY
Nonword repetition	SHORT-TERM MEMORY
Tonal sequence forward	SHORT-TERM MEMORY

Atonal sequence forward	SHORT-TERM MEMORY
Presentation of a sequence of 5 tones - recognition of one tone (tonal)	SHORT-TERM MEMORY
Presentation of a sequence of 5 tones - recognition of one tone (atonal)	SHORT-TERM MEMORY
Static matrix span	SHORT-TERM MEMORY
Syllable span	SHORT-TERM MEMORY
Recognition of consonants	SHORT-TERM MEMORY
Recognition of digits	SHORT-TERM MEMORY
Digit span backward	WORKING MEMORY
Reading span	WORKING MEMORY
Operation span	WORKING MEMORY
Test of Memory and Learning - Digits backward	WORKING MEMORY
Test of Memory and Learning - Letters backward	WORKING MEMORY
Spatial span backward	WORKING MEMORY
Two-back task	WORKING MEMORY
Presentation of a syllable and a sine wave tone simultaneously - tone recognition	WORKING MEMORY
Presentation of a syllable and a sine wave tone simultaneously - syllable recognition	WORKING MEMORY
Digit span forward with articulatory suppression	WORKING MEMORY
WAIS-IV letter-number sequencing	WORKING MEMORY
Visuospatial span	WORKING MEMORY

Note. Some tasks were present in more than one study.

Procedure

The preliminary dataset was composed of 37 studies and 99 tasks. For each task score, we calculated the variance and Hedges' *g*, which is a measure of the effect size adjusted for small groups (Borenstein, Hedges, Higgins, & Rothstein, 2009). The *g* values were interpreted according to the criteria suggested by Cohen (1988): small effect = 0.2 to 0.5; medium effect = 0.5 to 0.8; large effect > 0.8 (J. Cohen, 1988). The effect size was calculated starting from various values: from raw mean scores, standard deviations, and sample sizes of the group of musicians and the group of nonmusicians; from the value of F (Fisher) or t (Student's t-distribution) when the previous data were not reported in the study. If none of the above data was reported, or if there was missing data (e.g., the number of participants), the authors were contacted. For the present meta-analysis, eleven authors were contacted. Three of them could provide us with the missing data; the other eight studies were excluded from the analysis.

The final dataset for our meta-analysis included 29 studies and 75 tasks. If there were multiple measures (i.e., tasks) of the same construct (e.g., two different tasks investigating verbal working memory) the effect sizes of these multiple measures were combined. We did this for 15 studies, using the Borenstein method (Borenstein, Cooper, Hedges, & Valentine, 2009), with the *Mad* package (Del Re, Hoyt, 2010) of the *R* software (R Development Core Team, 2011). This method is important because it allows avoiding an overestimation of the effect size when there are multiple measures. To do this, it combines different effect sizes for dependent groups by taking into account also the correlation that might exist between two or more non-independent measures.

After having combined multiple measures, the final dataset included 53 tasks (used in the 29 studies) that were divided as follows: 14 tasks (10 studies) were assessing long-term memory, 20 tasks (16 studies) were assessing short-term memory, and 19 tasks (16 studies) were assessing working memory. We ran a separate meta-analysis for each of the three memory systems, with the *R* software and, specifically, using the *Metafor* package (Viechtbauer & Viechtbauer, 2015)

As we saw from literature, the difference between musicians and nonmusicians seems to vary depending on the type of stimuli presented (e.g., with tonal stimuli they perform always better, with visual stimuli results are mixed), therefore, we included this variable as a moderator in the meta-analyses. Stimuli were thus classified as: verbal (i.e., words, letters, and numbers, either read or heard); visuospatial (i.e., figures; spatial positions of figures); and tonal (i.e., musical tones; melodies). Although some studies distinguished between visual and spatial memory (e.g., Klauer & Zhao, 2004, for working memory) visual and spatial tasks were combined in the present meta-analysis. This was done because of the limited number of studies that were including only spatial stimuli or only visual stimuli. For each memory system, firstly we ran a random-effects model meta-analysis using the restricted maximum likelihood method (Viechtbauer, 2010). Summarized Hedges' g values were also estimated for each of the meta-analysis (i.e., for each memory system) using a Bayesian approach, with the *bayesmeta* package (Roever, 2015). Since this is the first meta-analysis that compares the memory of musicians and nonmusicians, less informative priors were used for our model parameters. Specifically, for μ we used a normal prior with a mean of 0 and a standard deviation of 10, whereas for τ we adopted a uniform prior of parameters 0 and 3. Both the maximum likelihood approach and the Bayesian approach led to the same conclusion (see Table 2).

Table 2. *Summary results of meta-analysis by memory system and estimation method.*

Memory System	Number of Tasks	Maximum Likelihood Approach	Bayesian Approach
		Hedges' g (95% Confidence Interval)	*Hedges' g* (95% Bayesian Credible Interval)
Long-Term Memory	14	.293 (.076 - .511)	.290 (.051 - .548)
Short-Term Memory	20	.569 (.408 - .730)	.567 (.400 - .744)
Working Memory	19	.565 (.328 - .802)	.564 (.309 - .827)

In the following step, we explored the heterogeneity across studies (i.e., differences across studies' results) using forest plots as graphical representation. The heterogeneity was examined using the *Q*-statistic (Higgins, Thompson, Deeks, & Altman, 2003), which is distributed like the chi-square under the null hypothesis; a significant chi value indicates the presence of heterogeneity across studies' results. We then estimated the magnitude of the heterogeneity with the *I2* index (i.e., the proportion of observed variance that reflects differences in effect sizes (Borenstein et al., 2009). A high *I2* value (i.e., *I2* > 75%, Higgins et al., 2003) suggest that results differ substantially across studies, and this could depend on several factors: for example, the studies could have measured different constructs or had different designs. In contrast, a low *I2* value (i.e., *I2* < 50%; Higgins et al., 2003) suggests that results across studies are similar, and this can therefore represent a true, generalizable effect.

As a further step, we also considered the presence of publication bias in each of the three meta-analyses (i.e., long-term, short-term, and working memory). Publication bias (Rothstein, Sutton, & Borenstein, 2006) is the phenomenon for which studies that report a statistically significant result (e.g., a difference between musicians and nonmusicians) are more likely to be published than studies that report a null result (e.g., no difference between musicians and nonmusicians). The publication bias was assessed using the funnel plot with the *trim and fill* method (Borenstein et al., 2009; Duval, 2006).

A technique to investigate the robustness of results is the sensitivity analysis. We used it in the present meta-analysis by applying the *leave one-out* method (Viechtbauer, 2010), which computes several meta-analyses on the dataset, leaving one study out each time. If the mean effect size changes substantially when a given study is removed, this means that the value of the mean effect size does not reflect the true mean (i.e., the mean depends largely on a single study), and that the studies lack homogeneity.

Finally, the role of the type of stimuli presented in the tasks was considered as a moderator, and examined using mixed-effects models (i.e., the type of stimuli was included as a fixed effect). The effect of the moderator (i.e., the type of stimuli) was tested using Wald's chi-square (Viechtbauer, 2010). Pairwise planned comparisons were used as well to explore the difference between the levels of the moderator (i.e., each level corresponded to a specific type of stimuli: visuospatial, verbal, and tonal). These comparisons were not orthogonal, therefore the type I error was controlled using the false discovery rate (Higgins et al., 2003). Table 3 shows the estimated means and 95% confidence intervals of the mixed-effects model.

Table 3. *Analysis of the moderating effect of the type of stimuli by memory system.*

Memory system	Tonal (*95% CI*)	Verbal (*95% CI*)	Visuospatial (*95% CI*)	Pairwise comparisons
Long-Term Memory	.01 (-1.03 – 1.04) n = 1	**.44** **(.16 – .73)** n = 8	.12 (-.22 – .45) n = 5	No difference
Short-Term Memory	**1.15** **(.79 – 1.51)** n = 4	**.54** **(.38 – .71)** n = 11	**.28** **(.04 – .52)** n = 5	Ton > Verb; Ton > Vis
Working Memory	**1.04** **(.48 – 1.60)** n = 3	**.59** **(.34 – .84)** n = 13	.01 (-.50 – .52) n = 3	Ton > Vis

Note. Estimated mean, 95% confidence intervals (*CI*) of summarized Hedges' *g*, and number of tasks by memory system and type of stimuli, calculated with the mixed-effects random models. Effect sizes significantly different from 0 at $p < .05$ are shown in bold. Significant pairwise differences between levels of the type of stimuli are displayed in the last column (i.e., pairwise comparisons). Ton = tonal; Verb = verbal; Vis = visuospatial.

Results

Descriptive statistics

The studies included in our meta-analysis were conducted between 1987 and 2017. The mean age of participants was 23.38 years (*SD* = 4.67). The samples of participants varied between 20 and 140 participants (*M* = 45.96, *SD* = 23.39), and they were always divided into two groups: musicians and nonmusicians. The studies included reported the duration of the musicians' music training in different ways: some of them reported the minimum years of music training, some others reported the mean years of music training, and some studies reported both the minimum and the mean. Across the studies that provided one of these two types of information, the minimum duration of music training was four years, while the average duration was 13.73 years. Table 4 shows the single effect sizes for each task included in the three meta-analyses.

Table 4. *Effect size and details of each task included in the three meta-analyses*

AUTHORS	PUBLICATION YEAR	MEMORY SYSTEM	TYPE OF STIMULI	*n* M	*n* NM	*g*	Var	Mean Age (*yrs*)
Amer, Kalender, Hasher, Trehub, & Wong.	2013	STM	VISUOSPATIAL	18	24	.87	.106	60
Anaya, Pisoni & Kronenberger	2016	STM	VERBAL	24	24	.52	.086	22.08
Bialystock & De Pape	2009	STM	VISUOSPATIAL	22	24	.42	.086	24.25
Bialystock & De Pape	2009	WM	VISUOSPATIAL	22	24	.39	.086	24.25
Boebinger & Evans	2015	STM	VERBAL	25	25	.19	.080	27.2
Boebinger & Evans	2015	WM	VERBAL	25	25	.30	.081	27.2
Brandler & Rammsayer	2003	LTM	VERBAL	35	35	.19	.044	28.45
Brandler & Rammsayer	2003	LTM	VISUOSPATIAL	35	35	-.06	.057	28.45
Chan, Ho, & Cheung	1998	LTM	VERBAL	30	30	.93	.056	19.75
Chan, Ho, & Cheung	1998	LTM	VISUOSPATIAL	30	30	.18	.050	19.75
Clayton et al.	2016	WM	VERBAL	17	17	1.01	.127	23.5
Franklin et al.	2008	LTM	VERBAL	12	13	.57	.119	19.73
Franklin et al.	2008	WM	VERBAL	11	9	.95	.170	21.6
George & Coch	2011	STM	VERBAL	16	16	.62	.098	20.25
George & Coch	2011	WM	VERBAL	16	16	.60	.098	20.25
George & Coch	2011	STM	VISUOSPATIAL	16	16	.56	.098	20.25
Hanna-Pladdy & Gajewski	2011	LTM	VERBAL	33	37	.35	.043	68.63
Hanna-Pladdy & Gajewski	2011	LTM	VISUOSPATIAL	33	37	-.14	.037	68.63
Hanna-Pladdy & Gajewski	2011	WM	VERBAL	33	37	.47	.058	68.63
Hansen, Wallentin, & Vuust	2012	STM	VERBAL	20	20	.97	.112	21.05

Hansen, Wallentin, & Vuust	2012	WM	VERBAL	20	20	-.06	.100	21.05
Hansen, Wallentin, & Vuust	2012	STM	VISUOSPATIAL	20	20	.42	.102	21.05
Hansen, Wallentin, & Vuust	2012	WM	VISUOSPATIAL	20	20	-.21	.101	21.05
Helmbold, Rammsayer & Altenmueller	2005	LTM	VERBAL	70	70	.06	.021	22.5
Helmbold, Rammsayer & Altenmueller	2005	LTM	VISUOSPATIAL	70	70	-.03	.029	22.5
Huang et al.	2010	LTM	VERBAL	10	10	.90	.220	21.45
Jakobson, Lewycky, Kilgour, & Stoesz	2008	LTM	VERBAL	15	21	.87	.083	19
Jakobson, Lewycky, Kilgour, & Stoesz	2008	LTM	VISUOSPATIAL	15	21	.82	.093	19
Lee, Lu, & Ko	2007	STM	VERBAL	20	20	.58	.078	22
Lee, Lu, & Ko	2007	WM	VERBAL	20	20	-.31	.077	22
Lee, Lu, & Ko	2007	WM	VISUOSPATIAL	20	20	-.17	.100	22
Monahan, Kendall, & Carterette	1987	STM	TONAL	12	10	1.02	.193	n.d.
Okhrey, Kutsenko, & Makarchuk	2017	STM	VERBAL	28	36	.29	.046	20
Okhrey, Kutsenko, & Makarchuk	2017	STM	VISUOSPATIAL	28	36	-.27	.062	20
Pallesen et al.	2010	STM	TONAL	11	10	1.42	.239	26.5
Pallesen et al.	2010	WM	TONAL	11	10	.51	.197	26.5
Parbery-Clark, Strait, Anderson, & Hittner	2011	WM	VERBAL	18	19	1.30	.126	50
Ramachandra, Meighan, & Gradzki	2012	STM	VERBAL	30	30	.78	.054	19.45
Ramachandra, Meighan, & Gradzki	2012	WM	VERBAL	30	30	.72	.053	19.45
Rodrigues, Loureiro, & Caramelli	2014	STM	VISUOSPATIAL	38	38	-.14	.040	32.15
Schiavo & Timmers	2016	LTM	TONAL	10	10	.01	.183	24.75
Schulze et al.	2011	WM	TONAL	16	17	1.44	.153	24.49
Schulze et al.	2011	WM	VERBAL	16	17	.43	.124	24.49
Schulze, Dowling, & Tillman	2012	STM	TONAL	20	20	.96	.084	22.68
Schulze, Dowling, & Tillman	2012	WM	TONAL	20	20	1.08	.086	22.49
Schulze, Mueller, & Koelsch	2011	STM	TONAL	16	17	1.37	.114	24.49
Suàrez, Elangovan, & Au	2016	WM	VERBAL	24	30	.62	.079	22.59
Suàrez, Elangovan, & Au	2016	STM	VISUOSPATIAL	24	30	.45	.058	22.59
Suàrez, Elangovan, & Au	2016	STM	VERBAL	24	30	.43	.077	22.59
Suàrez, Elangovan, & Au	2016	LTM	VERBAL	24	30	-.19	.075	22.59
Talamini, Carretti & Grassi	2016	STM	VERBAL	18	18	.66	.079	22.6
Talamini, Carretti & Grassi	2016	WM	VERBAL	18	18	.36	.075	22.6
Taylor & Dewhurst	2017	LTM	VERBAL	20	20	.66	.101	21.67
Vasuki, Sharma, Demuth, & Arciuli	2016	STM	VERBAL	17	18	.58	.114	25.75
Vasuki, Sharma, Demuth, &	2016	WM	VERBAL	17	18	.14	.128	25.75

Arciuli								
Weiss, Biron, Lieder, Granot, & Ahissar	2014	STM	VERBAL	42	15	.54	.093	23.35
Zuk, Benjamin, Kenyon, & Gaab.	2014	WM	VERBAL	15	15	1.19	.157	24.8

Note. The effect size is expressed by the g of Hedges. Additionally, for each task, information about the authors, the year of publication of the study, the memory system tapped, the type of stimuli presented in the memory task, the number of participants and the mean age of participants is reported. LTM = long-term memory; STM = short-term memory; WM = working memory; M = musicians; NM = nonmusicians.

Long-term memory

The random effect analysis revealed a small mean effect size, *g* = .29, 95% *CI* (.08±.51), *p* = .008, meaning that musicians tended to perform better than nonmusicians in long-term memory tasks, but the distance between groups was not large. The heterogeneity was significant, *χ2* (13) = 33.45, *p* = .001, *I2* = 63.71%, suggesting that the results of different studies vary moderately (see Figure 2). The sensitivity analysis indicated that the mean effect size did not vary consistently when removing single studies. In fact, the *g* value varied between .22 and .33 (*mean* = .29, *SD* = .03), meaning that the effect size remained small regardless of which study was excluded. We also assessed the publication bias, and in the funnel plot with trim and fill two hypothetical missing studies were added (see Figure 3). When including these two hypothetical studies in the meta-analysis, the effect size was reduced and became no longer significant, *g* = .21, 95% *CI* (-.02±.44), *p* = .068. The test of the moderator was not significant: *χ2* (2) = 2.42, *p* = .298 (for more details see, Table 2).

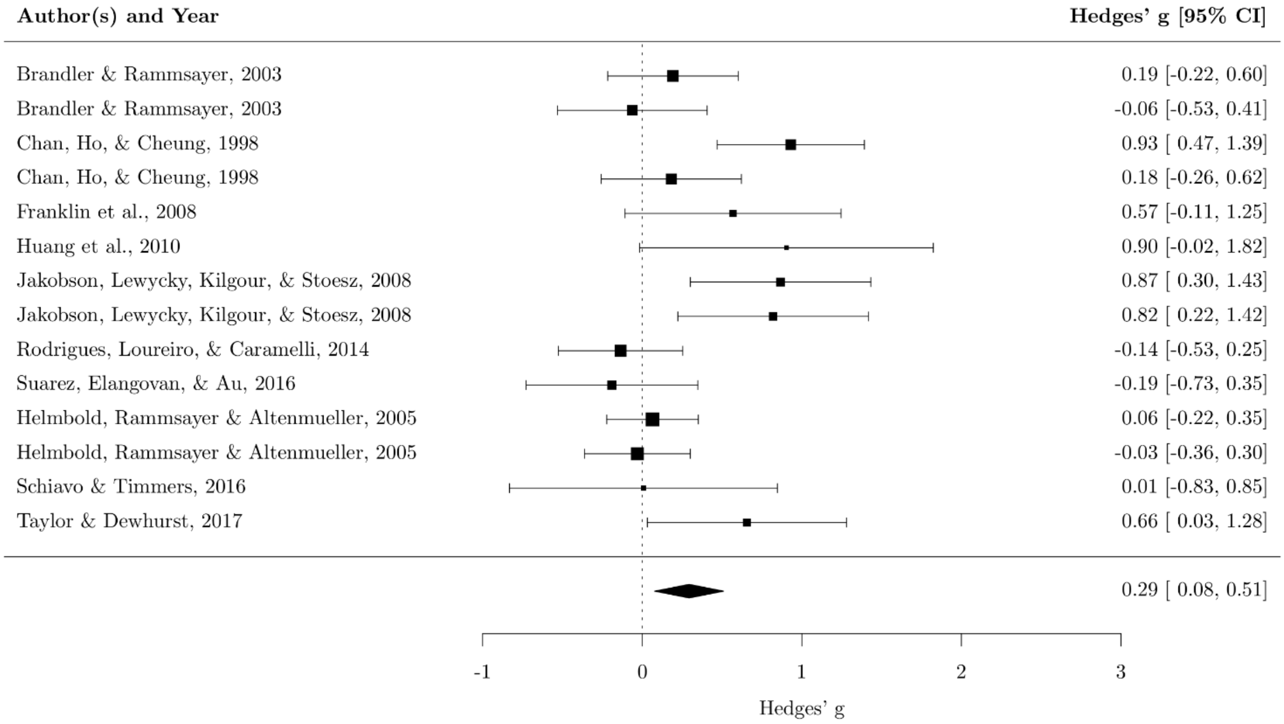

Figure 2. Forest plot for long-term memory. Each square represents the effect size of the study together with 95% confidence interval. The size of the symbol is proportional to the study's weight.

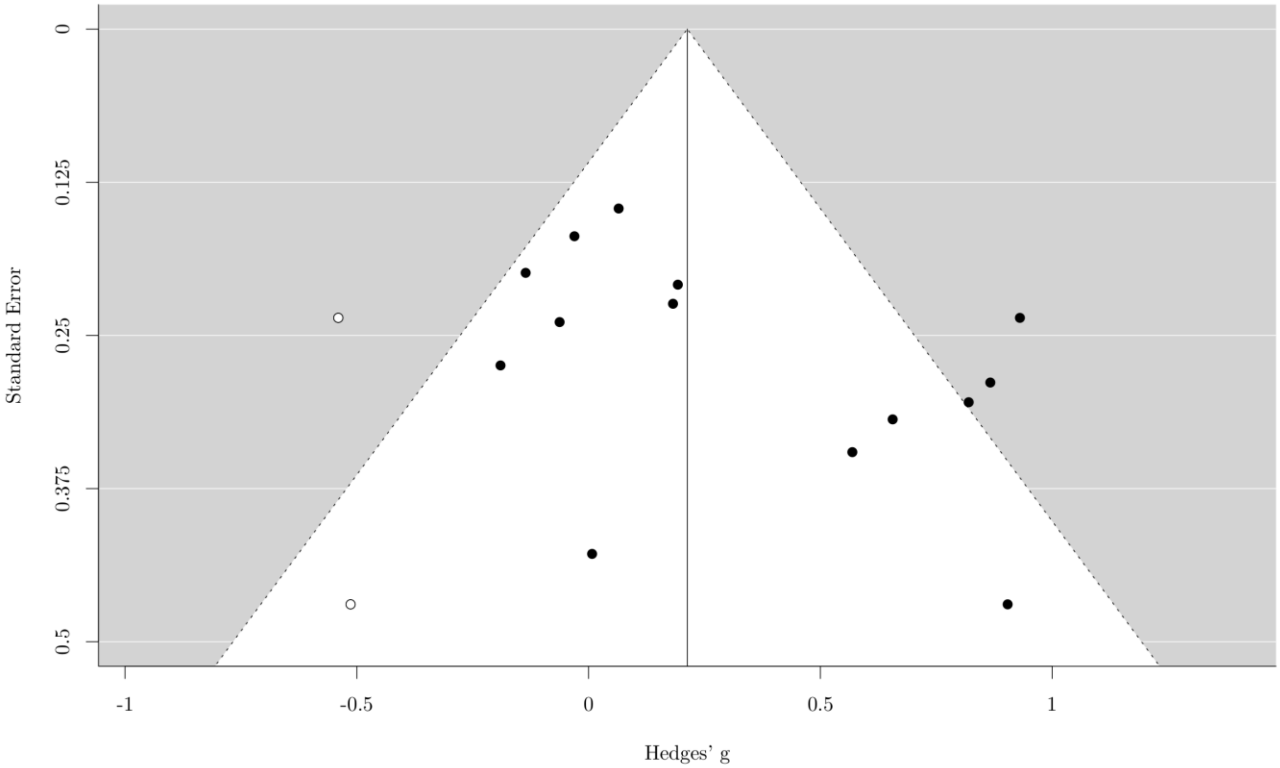

Figure 3. Funnel plot for long-term memory. Each black dot represents one study included in the meta-analysis. Any white dots represent the effect size of hypothetical unpublished results.

Short-term memory

The random effect analysis revealed a moderate mean effect size, $g = .57$, 95% *CI* (.41±.73), $p < .001$, indicating that musicians performed better than nonmusicians in short-term memory tasks included. The heterogeneity was not significant, $\chi 2$ (19) = 29.67, $p = .056$, $I2 = 35.36\%$, meaning that most of the studies had similar results (see Figure 4). The sensitivity analysis showed that the mean effect size was robust even when removing single studies. In fact, the *g* value ranged between .53 and .61 (*mean* = .57, *SD* = .02), depending on which study was excluded. When assessing publication bias, the funnel plot with trim and fill added seven hypothetical missing studies (see Figure 5), but even when including these hypothetical studies in the analysis the effect size did not change substantially, and remained moderate, $g = .39$, 95% *CI* (.21±.57), $p < .001$.

Even if the heterogeneity was not significant (i.e., there was low variance across studies' results), we investigated whether the moderator (i.e., the type of stimuli) could influence the effect size. The test on the moderator was significant, $\chi 2$ (2) = 15.64, $p < .001$, and the heterogeneity was still not significant, $\chi 2$ (17) = 14.02, p = .66, $I2$ = 0.04%. When adding the moderator in the meta-analysis, the amount of heterogeneity decreased, suggesting that the type of stimuli played a role in determining the small differences across studies' results. Specifically, for all the levels of the moderator, the effect size was statistically different from zero, with tonal stimuli associated to the largest effect size, and verbal and visuospatial stimuli associated to a moderate effect size (see Table 2).

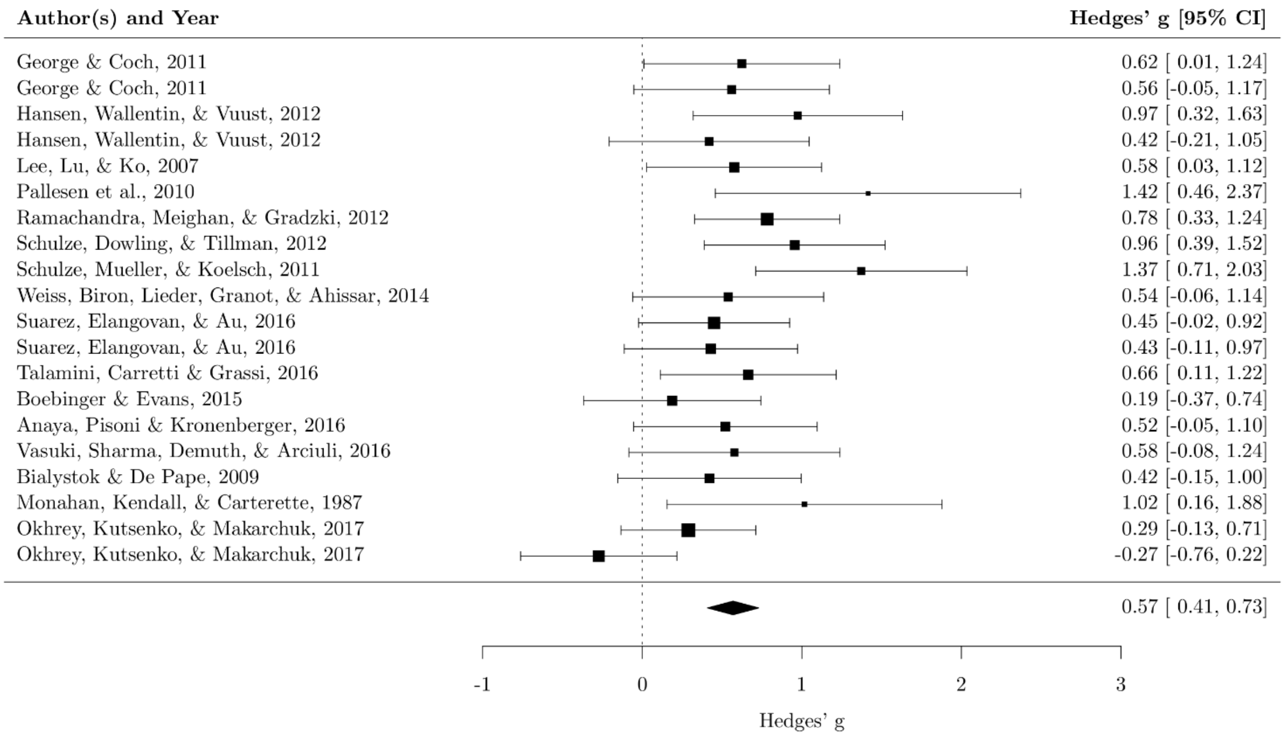

Figure 4. Forest plot for short-term memory. Each square represents the effect size of the study together with the 95% confidence interval. The size of the symbol is proportional to the study's weight.

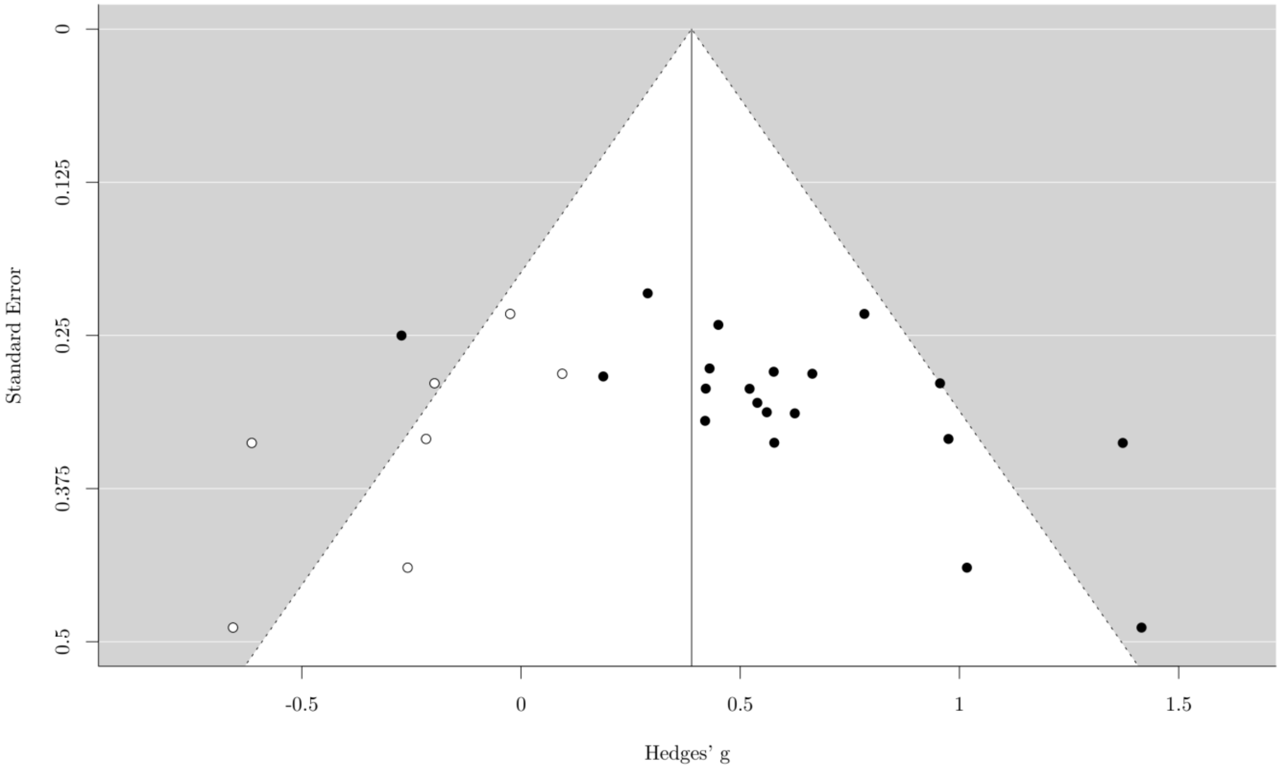

Figure 5. Funnel plot for short-term memory. Each black dot represents one study included in the meta-analysis. Any white dots represent the effect size of hypothetical unpublished results.

Working memory

The random effect analysis revealed a moderate mean effect size, *g* = .56, 95% *CI* (.33±.80), *p* < .001, meaning that, also in working memory tasks, musicians performed better than nonmusicians. In this case, the test of heterogeneity was also significant, *χ2* (18) = 47.41, p < .001, *I2* = 62.85%, revealing a moderate variance across studies' results (see Figure 6). The sensitivity analysis, by removing single studies, indicated that the mean effect size was robust. In fact, the *g* value varied from .52 to .62 (*mean* = .56, *SD* = .03), showing that none of the studies included in the meta-analysis had a substantial influence on the mean effect size. The funnel plot with trim and fill did not show any evidence of publication bias (see Figure 7).

The analysis of the moderator was significant, *χ2* (2) = 7.36, *p* = .025. However, the test for residual heterogeneity was still significant, *χ2* (16) = 32.73, *p* = .008, *I2* = 51.42%, suggesting that the moderator could explain only partly the variance across studies' results, and therefore the moderator is not sufficient to explain this variability observed. Specifically, two of the three levels of the moderator had an associated effect size significantly different from zero: tonal stimuli were associated with the largest effect size, followed by verbal stimuli, that were associated to a moderate effect size. No effect was found (i.e., no difference between musicians and nonmusicians) for visuospatial stimuli (see Table 2 for details).

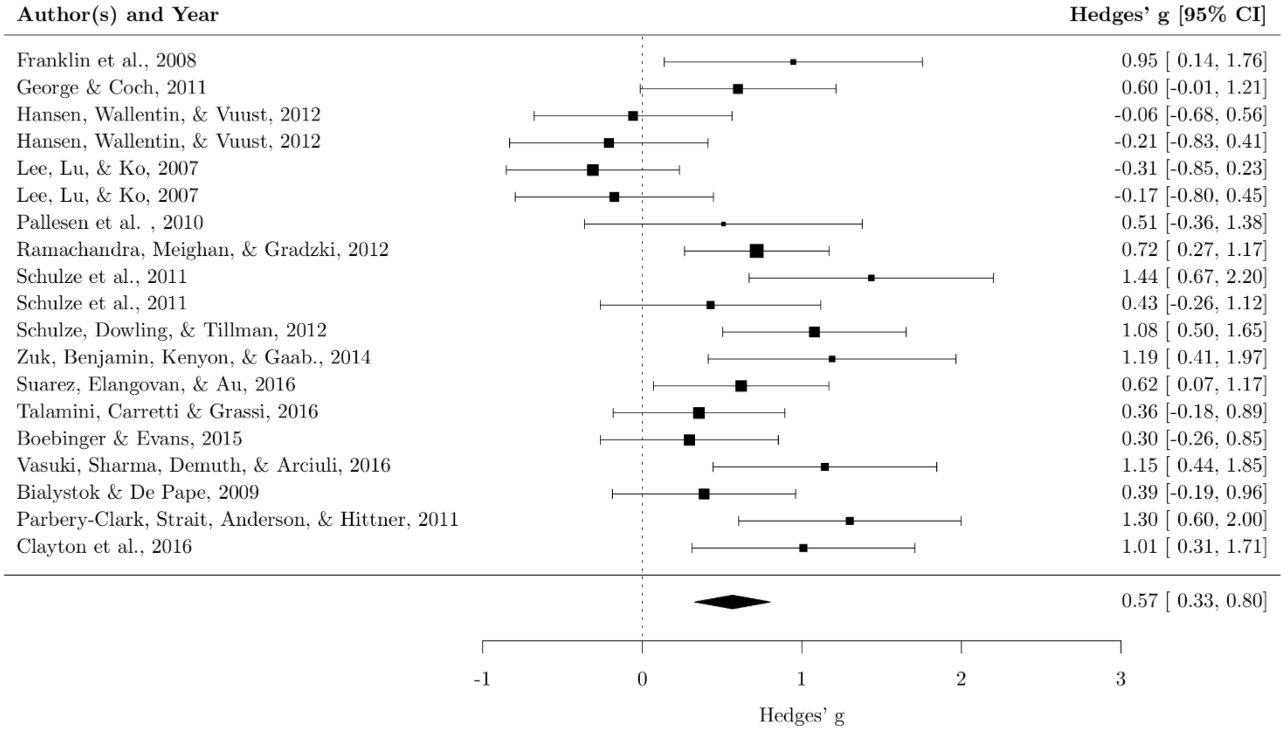

Figure 6. Forest plot for working memory. Each square represents the effect size of the study together with the 95% confidence interval. The size of the symbol is proportional to the study's weight.

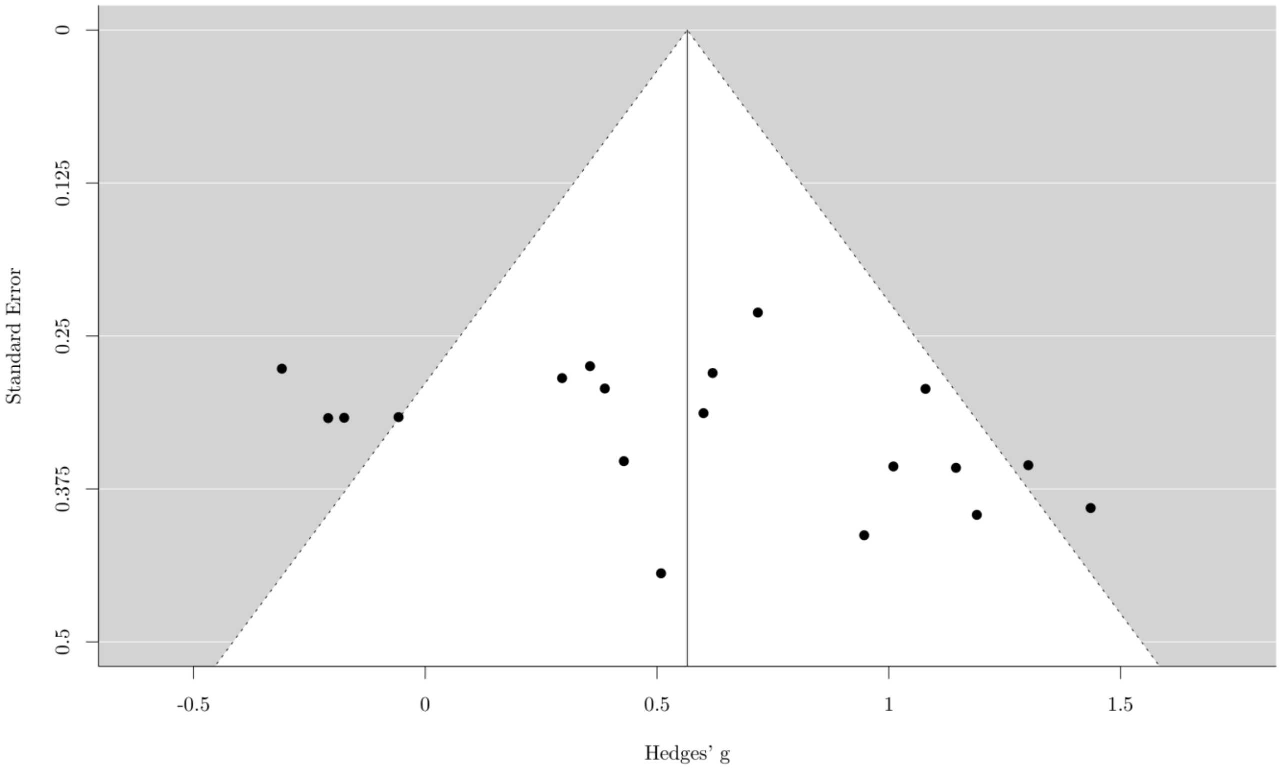

Figure 7. Funnel plot for working memory. Each black dot represents one study included in the meta-analysis. Any white dots represent the effect size of hypothetical unpublished results.

Discussion

With the present meta-analysis, we aimed to investigate whether musicians have better memory than nonmusicians, considering separately long-term memory, short-term memory, and working memory. We wanted also to examine whether a possible advantage of musicians over nonmusicians in memory performance could be modulated by the type of stimuli presented in the tasks (i.e., verbal, visuospatial, and tonal). As emerged from the literature reviewed in Chapter 2 and 3, musicians often perform better than nonmusicians in various cognitive domains (including memory). According to the literature, it could be hypothesized that the performance of musicians in memory tasks is enhanced for domain-specific stimuli (e.g., tones), with which they are familiar (e.g., Ericsson & Charness, 1994). The meta-analysis was thus conducted considering memory tasks that tap different memory systems, and that use different types of stimuli. Overall, the present findings support a domain-specific superiority of musicians over nonmusicians in memory tasks. Nevertheless, the domain specificity hypothesis is not enough to explain all the reported results, which will be discussed below, separately for memory system and type of stimuli.

Long-term memory. The meta-analysis revealed a small advantage of musicians over nonmusicians, with a moderate variability across results of different studies. The funnel plot showed two hypothetical unpublished studies with null or opposite results (i.e., nonmusicians performed better than musicians), suggesting that a possible publication bias exist. Moreover, when these two hypothetical studies were included in the analysis, the difference between musicians and nonmusicians decreased and became no longer significant. Therefore, further studies on long-term memory are needed to clarify whether the difference previously identified between musicians and nonmusicians reflects a true effect or not. In addition, the moderator (i.e., the type of stimuli) was unable to explain the heterogeneity across studies' results, which decreased only slightly when the moderator was included in the analysis.

Concerning the type of stimuli, it should be noted that, differently from the studies on short-term and working memory, here, only one study included in the meta-analysis investigated the recall of tonal stimuli in long-term memory (i.e., Schiavio & Timmers, 2016). In this study participants had to learn and remember short ambiguous melodies, and the authors observed no difference between musicians and nonmusicians. On the contrary, studies using musical stimuli in test short-term tasks and working memory tasks, observed differences between groups, and with large effect sizes. The fact that, in long term memory, we had only one study that used musical stimuli might be one of the reasons behind the null effect of the moderator in the meta-analysis. Another reason of this null effect could be due to the difference between the methodologies used in the studies. Nevertheless, although the test of moderator was not significant, we observed that tasks that presented verbal stimuli were associated with a larger effect size than tasks that used visuospatial stimuli. This result in line with the result observed for working memory.

Short-term memory. The meta-analysis revealed a moderate effect size, suggesting that musicians outperform nonmusicians in short-term memory tasks. Moreover, there was no significant heterogeneity across studies. As in long-term memory, the funnel plot suggested a publication bias. Nevertheless, when adding the hypothetical missing studies, the effect size did not change substantially, and remained moderate, strengthening the reliability of the result of the meta-analysis for short-term memory. In this case, the moderator analysis revealed a significant effect of the type of stimuli. Specifically, when the moderator was included in the analysis, the heterogeneity, that was already statistically insignificant, almost disappeared completely. In other words, in short-term memory tasks, the musicians' advantage changed depending on the type of stimuli presented in the task: as suggested by the domain specificity hypothesis, musicians outperformed nonmusicians especially when the task asked to remember music stimuli. It is worth noting, however, that they still performed better than nonmusicians with also verbal and visuospatial stimuli.

Working memory. Here, the meta-analysis showed a moderate effect size, indicating, again, that musicians outperformed nonmusicians in the working memory tasks considered. Nevertheless, the results across studies revealed a moderate variability: this means that some studies reported a superiority of musicians over nonmusicians, whereas others did not (or not as much). When the moderator was included in the analysis, the heterogeneity remained significant, but decreased slightly, suggesting that the type of stimuli presented in the tasks can explain only a small part of the variability observed across studies. Specifically, tonal stimuli were associated with the largest effect size, again supporting a domain specific advantage. Verbal stimuli were also associated with a moderate effect size, but, interestingly, with visuospatial stimuli the mean effect size was not significant, meaning that there was no difference between groups. Furthermore, the funnel plot suggested that there is no publication bias.

To sum up, the present meta-analysis suggests that musicians have better memory than nonmusicians. We might ask ourselves, whether the difference between groups observed is genuine or if it depends on the journals' policy to publish positive (rather than null) results (i.e., the publication bias, see Munafò et al., 2017) . Of course, in the case that the currently-available literature only contains positive results, then the outcome of our meta-analysis would necessarily reflect a publication bias. By the same token, if this were true, the statistical methods that we adopted to check for a publication bias would be of little usefulness. While we cannot exclude that such a bias exists, some of the results obtained in the present meta-analysis support the idea that this effect is, to same extent, real (i.e., musicians really do have better memory than nonmusicians). For example, we observed a difference between the mean effect sizes for long-term memory (small) and the one for short-term and working memory (moderate). Moreover, the effect of the moderator suggests that the results we observed reflect a real difference, at least in part, because the musicians' advantage was not constant, but it was large for tonal stimuli, medium for verbal stimuli, and small-to-null for visuospatial stimuli.

That said, the present meta-analysis showed that musicians perform better than nonmusicians in memory tasks; this raises a question: why musicians have better memory than nonmusicians? If considering tonal stimuli, the advantage here observed can be easily explained according to the literature on experts' performance. In fact, experts (musicians in our case) perform better than non-experts (i.e., nonmusicians) with tasks and/or with stimuli they are familiar with. What cannot be explained is why musicians perform better than nonmusicians in memory tasks with verbal stimuli and (to some extent) with visuospatial stimuli too. We can hypothesize two possible explanations for this extended advantage. The first one is that there might be some uncontrolled variables, specific to quasi experiments, that are responsible for the difference between musicians and nonmusicians in memory performance. One example can be that musicians perform better than nonmusicians because of a sort of Pygmalion (or Rosenthal) effect (Rosenthal & Jacobson, 1968). If researchers expect musicians to perform better, this could induce an improvement in their performance. Nevertheless, this possibility will not explain why a difference between musicians and nonmusicians is evident at different degrees for memory, and it is also absent in some cases with other cognitive abilities (see for a broad overview Schellenberg & Weiss, 2013). Another factor that can explain these results that quasi experiments cannot totally control, is that individuals with better memory are more likely to become musicians, therefore there was already a difference between groups before the musicians started a music training. The same issue concerns also any individual characteristic that might help a participant to do well in memory tasks (e.g., enhanced sensory abilities, intelligence, personality, etc.) (Corrigall, Schellenberg, & Misura, 2013). On the other hand, any of these possible factors, if responsible of the advantage of musicians, would give them a constant superiority across the various memory systems and types of stimuli presented in the task, but this did not emerge in the present meta-analysis, and, on the contrary, the advantage varied consistently depending on what was tested.

Moving away from all the possible factors contributing to this advantage that are not controlled in quasi experiments, a second possible explanation of these results is that a better memory might be a consequence of having undergone a music training. Learning to play an instrument might improve the individual's memory for tonal stimuli (i.e., the domain specificity hypothesis, Ericsson & Kintsch, 1995)) and this would explain why musicians perform better than nonmusicians in memory tasks that involves tonal stimuli. However, the present findings indicate that musicians’ advantage extends also to verbal stimuli. As mentioned in chapter 2, musicians process auditory stimuli better than nonmusicians (e.g., Rammsayer & Altenmüller, 2006; Spiegel & Watson, 1984; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). We can thus hypothesize that, thank to this ability, musicians perform better in memory tasks when stimuli are presented orally, because they might have a better auditory encoding of the item to be remembered, that could strengthen the memory trace of the stimulus. This might explain why musicians perform better than nonmusicians with verbal material too (the stimulus modality hypothesis). In memory tasks, verbal stimuli (e.g., words, numbers, syllables, etc.) are often presented orally, like in most of the studies considered for the present meta-analysis). This hypothesis is supported, for example, in the study by Okhrei and colleagues (2017), who found no difference between musicians and nonmusicians in short-term memory when verbal stimuli were presented visually in the task (Okhrei et al., 2017). Similar results were reported by our first experiment (see STUDY 2) in which musicians performed better than nonmusicians in the digit span task when the digits were presented orally, but the difference between groups was smaller (and nonsignificant, in the second analysis) when they were presented visually (Talamini et al., 2016). Another possible explanation for the advantage of musicians in memory tasks with verbal stimuli concerns the relationship between music and language: some authors support the idea of shared processes between music and language (see Anvari et al., 2002; Patel, 2003 for an overview). Music perception skills are related to phonological awareness and early reading development (e.g., Patel & Iversen, 2007). Some music perception skills (for a description of music aptitude see chapter 3), for instance, predict reading skills, suggesting that some music perception skills are related to auditory or cognitive mechanisms that are used also when reading (Anvari et al., 2002). However, these hypotheses cannot explain why musicians sometimes are found to have an advantage with visuospatial stimuli too.

Finally, the superior performance of musicians might be related to the multisensorial nature of the music training. In fact, learning to play a musical instrument involves associating the various symbols of music notation with the sound of the notes, and the appropriate motor response (which has to be address to a specific spatial position). The person first learns to associate the music notation with sounds and motor actions, through specific exercises. This training requires initially a large effort, and it demands attentional control. After a while, however, there is a decrease in the attentional control needed over the learning process, because the individual learns to associate notes, sounds and actions more automatically (see chapter 1). Becoming an expert, in this case a musician, will be therefore linked to enhanced active and controlled learning skills, which could be helpful when remembering any kind of stimuli (i.e., visuospatial too) in memory tasks. In other words, music training might benefit active learning strategies, such as, for example, chunking strategies. In fact, when learning a music score, chunking is essential in order to be able memorizing an entire melody. Chunking improves the capacity to memorize series of items, and it might be that musicians perform better than nonmusicians in short-term and working memory tasks because they are able to use more efficient chunking strategies. In any case, the best way to investigate the possible reason of the advantage of musicians over nonmusicians in memory tasks would be by running appropriate experiments, such as longitudinal studies with participants randomly assigned to different groups (e.g., a music training group and other groups with other training activities). Unfortunately, the present results do not allow drawing sure conclusions.

Meta-analyses are powerful and useful techniques without doubts, however, it is important to underline some limits of the present work. The first limit regards the number of studies included in the analysis, 29. These studies were also divided into three groups, leading with a relatively small number for each memory system. Moreover, each moderator level (visuospatial, tonal, verbal stimuli), so some levels of the moderator (i.e., visuospatial, tonal, and verbal stimuli) were under-represented (e.g., tonal stimuli in long-term memory). For this reason, even if the effect sizes found support a difference between musicians and nonmusicians in several tasks, these results should be interpreted with caution.

A second limitations of the current study is that the years of music training received by the musicians' groups could not be controlled, because studies reported this information in different ways, not always comparable from one another. Specifically, some studies reported the mean of the total years of music training of the musicians' group; others, instead, reported the minimum number of years of music training. As a consequence, we could not include this variable in the meta-analysis, and it might have been important in explaining part of the heterogeneity observed across studies. Moreover, as already mentioned in the limitations of literature in chapter 3, there is currently no standard for describing and including musicians and nonmusicians in these experiments, (e.g., the minimum of years of training that musicians should have undergone, and how to report this information). Moreover, several potentially interesting characteristics are very often not reported (e.g., hours of daily practice, instrument played, etc.). In literature, there are some questionnaires that can be used to have a complete description of participants (both musicians and nonmusicians, (Chin & Rickard, 2012; Müllensiefen, Gingras, Musil, & Stewart, 2014) and, especially with quasi experiments, a detailed description of the two groups would be of fundamental importance, to exclude possible confounding variables. Often, having too little information does not help disentangling whether musicians' enhanced performance is an effect of their training. The lack of control variables, that might explain the difference between groups, is also an important issue, because several studies did not include them: for instance, not all the studies considered here controlled for general cognitive abilities (e.g., intelligence).

Despite the limitations of the present meta-analysis, this work can have several advantages. In fact, it can help underlining the weaknesses of past studies that compared musicians and nonmusicians, making it easier for future research to overcome these limitations. As already mentioned, longitudinal studies on music training would be ideal to investigate the possible effects of the music training on cognitive abilities. In the case, instead, of quasi-experimental studies that compare musicians and nonmusicians should provide as many details as possible about their participants and they should also control for general cognitive abilities, socio-economic status, and personality, in order to take into account potential pre-existing differences.

To conclude, this meta-analysis revealed that musicians perform better than nonmusicians in several memory tasks. We have listed several possible explanations for this advantage, but, at the current state of literature, none of them seem able to explain all the results. It is likely that more than one mechanism lies behind the musicians' advantage, and that this advantage is partly domain-specific (for tonal stimuli). It is also possible that their advantage extends also to verbal memory tasks presented orally, because of the musicians' enhanced auditory perception. As previously mentioned, these explanations are not yet sufficient, and only further studies can reveal more about the musicians’ advantage in memory tasks.

Study 2. The Working Memory of Musicians and Nonmusicians

Authors: Francesca Talamini*, Barbara Carretti*, Massimo Grassi*

*Department of General Psychology, University of Padua, Italy

Published on Music Perception in March 2016, doi: 10.1525/MP.2016.34.2.183

Aims

This study aims to investigate verbal working memory of musicians, since literature suggested that musicians outperform nonmusicians in these tasks. It is interesting to explore this advantage because several authors claimed that there is a far transfer of the music-trained skills on verbal skills. However, a problem of the studies which focused on verbal memory, is that the tasks used to assess verbal STM and WM are mainly administered auditorily (e.g., the experimenter reads aloud the stimuli to the participant). The first objection was that, perhaps, the advantage of musicians over nonmusicians is not linked to verbal material per se, but it is driven by a general better auditory STM and WM (e.g., independently from the kind of auditory material used). The aim of the current study was to assess this hypothesis by administering a verbal STM and WM tasks (i.e., a digit span with and without articulatory suppression) not only auditorily, but also visually, and audiovisually. We expected musicians to perform better only in the auditory condition if their advantage was driven by a better auditory STM and WM, and not by generally enhanced verbal skills. In contrast, if musicians performed better regardless of the modality of presentation, this could suggest that their advantage truly extends to verbal stimuli. In addition, we decided to assess also music aptitude of all participants, because only a few studies (e.g., Hansen et al., 2013; Wallentin et al., 2010) did that before. We aimed to understand whether the music aptitude was correlated to the span performance, independently from having undergone a music training. Finally, we administered some tests to control for general cognitive abilities, because it allows controlling whether a possible superiority in the memory task is driven by a general cognitive advantage of the participant.

Method

Participants. Thirty-six Italian young adults participated to the study. Eighteen participants were musicians, that is, they were all students of a music conservatory, and they had undergone a minimum of seven years of music training. The other 18 participants were defined as “nonmusicians”, because they had not received any music training apart from the general music classes that Italian students attend at the middle school. All participants were university students. Demographic details are reported in Table 1. A set of independent sample t-tests were calculated to understand whether musicians and nonmusicians differed in age, educational level, Vocabulary and Visual Puzzle scores. The results showed that nonmusicians had a higher educational level than musicians, $t(34)=3.17$, $p = .0032$. In contrast, the two groups did not differ in terms of age $t(34)=1.21$, $p = .24$, in the visual puzzle test, $t(34)=.5$, $p = .61$, and in vocabulary test, $t(34)=1.11$, $p = .27$ (see Table 1).

All participants were not familiar with psychological experiments. They had normal audition (assessed with an audiometry, for frequencies of 500, 1500, and 4000 Hz) and normal to corrected vision.

Table 1. *Age, Education, Performance (Raw Scores) in the WAIS-IV Visual Puzzles and Vocabulary Subtests*

	Musicians		Nonmusicians	
	M	*SD*	*M*	*SD*
Age (years)	22.06	3.80	23.28	2.02
Education (years)	14.94	2.15	16.78	1.17
Visual Puzzle (max score 26)	17.39	5.37	18.17	3.75
Vocabulary (max score 57)	44.72	7.75	42.17	5.82

Material

Digit span test. Digits were presented in three different modalities: visually, auditorily, and audiovisually. When presented visually, they appeared at the centre of the screen, for 750ms, one at time (digits were written in Arabic numbers; 24 points Arial font); when presented auditorily, they were delivered through a pair of headphones; finally, when presented audiovisually, they were showed on the screen and played through the headphones simultaneously. The subject was completing a block of trials always with the same modality of presentation (e.g., s/he could start with the auditory presentation, then, once finished, s/he was moving to visual one, and so on). The order of presentation of the three modalities was counterbalanced across subjects. The digits were randomly presented from 1 to 9, starting with a 3 digits sequence. There were two trials with the same digit length, and, as long as the participant could correctly reproduce at least one of the two sequences in the same order of presentation, the length of the sequence increased by one number. The participant responded by typing the sequence previously memorized on a computer keyboard. A pure tone of 500 Hz was presented for 500 ms before and after each sequence, to signal the begin and the end of it. In the visual modality, an asterisk was presented for 500 ms before and after the sequence. In all modalities, the numbers were presented at a pace of one every 1.5 seconds. In the auditory presentation, the pace was calculated as the temporal distance between the tonic accents of two consecutive numbers (e.g., the distance between /u/ and /e/ in the ‘‘uno – tre’’ sequence, respectively one and three in English). In the visual version, the pace was calculated as the temporal distance between the onsets of two consecutive numbers. The Digit span was administered with two different conditions, without and with articulatory suppression. In the articulatory suppression condition, the participants were asked to say aloud the syllables “la la la” while hearing/seeing the digits. This technique limits the use of rehearsal strategies, and cognitive load of the task.

Mini PROMS (Law & Zentner, 2012). This online test was used to assess the music aptitude of all participants. It consists of four subtests (Melody, Tuning, Beat and Speed) that investigate different aspect of the music perception. In each subtest, the participant listened twice to a standard stimulus that was followed by a comparison stimulus. The participant had to judge whether the comparison stimulus was identical or different from the standards. In half of the trials the comparison was identical to the standards, in the other half it was different. The answer was given on a five points scale, and the possible answers were the follows: “definitely same”, “probably same”, “I don’t know”, “probably different”. When the right answer was given with the maximum of confidence (i.e., “definitely same/different”), the score was two points. When the right answer was given with less confidence (i.e., “probably same/different), the score was one point. Finally, in the other three cases (i.e., “I don’t know” and both “definitely...” and “probably” that go in the wrong direction) the score was zero.

The Melody subtest assesses the ability of recognizing whether two different short melodies (the standard – played twice, and the comparison) are identical or not. The difficulty varies in terms of number of notes that composes the melody and atonality. There could be one note only or multiple notes that change in the comparison melody (when different). The Tuning subtest requires to compare chords. The chords are composed by four different notes (i.e., C4, E4, G4, and C5), and they last for 1.5. In the case of different trials, one of the middle notes of the comparison chord is different in frequency (i.e., it is mistuned). This note can be shifted from 10 cents to 50 cents, depending on the level of difficulty (i.e., the higher the shift, the easiest the trial). The Beat subtest requires to compare rhythmic patterns of clicks: the rhythm is produced by giving an accent (i.e., by changing the intensity, an increment of 3 dB) to a subset of the clicks. In easy trials the change in intensity is distributed among several clicks. In difficult trials it is distributed in few clicks only. In the Speed subtest the participant compares the speed (i.e., beats per minute, BPM) of either a synthetic rhythmic structure, or a recorded sample of music. The comparison stimulus, when different, has a different speed that is varied from þ1 BPM to þ7 BPM.

WAIS-IV (Wechsler, 2008): Puzzle subtest. This subtest investigates nonverbal reasoning. It requires to look (in a book placed in front of the participant) at a figure of reference and then choose three elements (in a pool of five) that can be combined together to recreate that figure. The participant can mentally rotate the various elements. The test is timed, and the participant has to give the answer in 20 sec or 30 sec (depending on the level of difficulty). The difficulty is increasing trial by trial, and after three consecutive trials with a wrong answer, the test is interrupted. Each correct answer scores 1 point.

WAIS-IV: Vocabulary subtest. This subtest is used to investigate verbal skills. The experimenter reads some words (one at time) to the participant, who has to give a short definition of the word presented. Complete definitions score 2 points, incomplete definitions score 1 point, wrong definitions score 0 points. The level of difficulty is increasing word by word, and after three consecutive scores of 0, the test is interrupted.

Apparatus

The computer used for the digit span test and the PROMS test was an ASUS (Cpu Intel i5 650 3.20 GHz, Motherboard Asus P7H55-V RAM 4 GB, Graphic Card AMD Radeon HD 5700 Series, OS Windows 7 Professional 64 bit). The computer was connected to a monitor (NEC MultiSync FE950þ) and M-AUDIO FastTrack Pro sound card. The headphones were a pair of Sennheiser HD 580. The computer tests were delivered inside of a single walled IAC sound proof booth. The digit span test was created with the software MATLAB, using a custom-coded program with the extension “Psychophysics Toolbox” (Kleiner et al., 2007). The auditory version of the digits in the span, were recorded by a male speaker with a neutral prosody, with a Shure SM 58 microphone. The editing of the recordings was done with the software Cool Edit Pro (Syntrillium Software), and then they were assembled with a MATLAB custom-coded program. The PROMS test was administered on its website.

Procedure

Participants began by signing a consent form, then information about demographic details (e.g., age, sex, education) was collected. Participants then sat into the soundproof boot, where the audiometry test was run. Successively, the first three blocks of the digit span (auditory, visual, and audiovisual) were administered in a counterbalanced way across subjects, and within groups (e.g., one participant started with the auditory modality of presentation, another one with the visual modality, and so on). Half of subjects (of each group) started with the no articulatory suppression condition, whereas the other half started with the articulatory suppression condition. After the first three blocks of digit span were administered, the experimenter entered in the boot, and administered the two WAIS-IV subtests, the Visual Puzzles and then the Vocabulary. After these two tasks, the participant completed the second three blocks of the digit span (in the other condition, with or without articulatory suppression, depending with what they started). After the second block of the digit span, participants completed the PROMS test. Finally, a questionnaire about music habits was administered. This questionnaire contained some questions about music habits for all participants (e.g., listening to music, dancing) and specific question for music education for the musician group (e.g., years of training, hours of practice, type of instrument).

Results

A two-ways ANOVA was run, with the span measure (i.e., the total number of digit sequences correctly reproduced) as dependent variable. There were two within-subjects factors, namely the modality of presentation (auditory, visual, and audiovisual), and the condition (with and without articulatory suppression); there was one between factor, that is, the group (musicians and nonmusicians). The ANOVA revealed an overall advantage of musicians over nonmusicians in the digit span, regardless of the modality and the condition, $F(1, 34) = 4.41$, $p = .04$, $\eta^2 = .12$. The condition was also significant, $F(1, 34) = 94.22$, $p < .001$, $\eta^2 = .74$, meaning that the performance of all participants was worse in the articulatory suppression condition. The condition did not interact with the group factor, $F(1, 34) = 1.98$, $p = .17$, $\eta^2 = .06$. The modality factor was close to be significant, $F(2, 68)$

= 2.81, *p* = .07, η^2 =.08, but again, there was no interaction with the group, *F* < 1. There was a significant interaction between suppression and modality, *F*(2, 68) = 4.09, *p* = .02, η^2 =.11, but no interaction with the group, *F*(2, 68) =1.47, *p* = .24, η^2 = .04. See Figure 1 for the span results.

Even if the ANOVA did not show any interaction between the group and the modality, we decided to further explore it because qualitatively it was possible to observe a tendency (i.e., musicians' advantage looked larger in the audio and audiovisual modality). We then ran three separate ANOVAs separately for the three different modalities of presentation. As between factor we had always the group, and as within factor the condition (i.e., articulatory suppression).

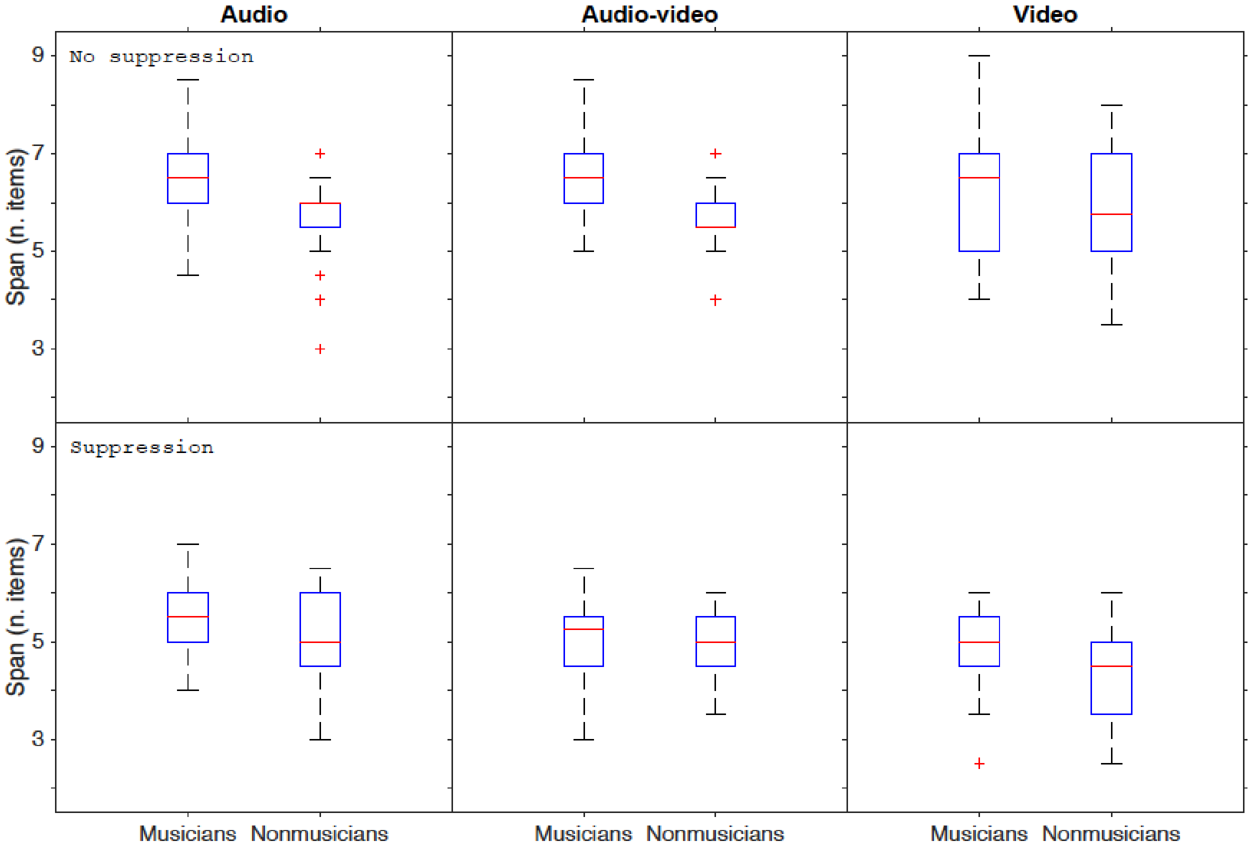

Figure 1. Performance of digit span in the three different modalities (i.e., auditory, audiovisual, and visual) and the two conditions (i.e., with and without articulatory suppression). In each box, the red horizontal line indicates the median. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The red crosses are the outliers.

AUDITORY PRESENTATION: The span presented in the auditory modality was larger for musicians than for nonmusicians, *F*(1, 34) = 5.06, *p* = .03, η^2 = .13. The dimension of the effect was large (d =.80) following Cohen's (1988) guidelines. The condition was again significant, meaning that, both groups performed better without articulatory suppression than with articulatory suppression: $F(1, 34) = 28.68$, $p < .001$, $\eta^2 = .46$. There was no interaction between group and condition, $F(1, 34) = 1.14$, $p = .29$, $\eta^2 = .03$.

AUDIOVISUAL PRESENTATION: Musicians performed better than nonmusicians when the span was presented audiovisually, $F(1, 34) = 4.19$, $p = .02$, $\eta^2 = .11$. The condition was again significant, meaning that both groups performed better when there was no articulatory suppression, $F(1, 34) = 48.06$, $p < .001$, $\eta^2 = .59$. In addition, the interaction between group and condition was significant, $F(1, 34) = 6.58$, $p = .02$, $\eta^2 = .17$. Post hoc analyses showed that musicians had an advantage over nonmusicians in the audiovisual modality with no articulatory suppression ($p = .005$), with a large effect size ($d = 1.01$), but there was no difference between groups in the articulatory suppression condition ($p = .63$).

VISUAL PRESENTATION: Finally, when the span was presented visually, there was no difference between the two groups, $F(1, 34) = 1.35$, $p = .25$, $\eta^2 = .04$. Condition was always significant, with larger span for both groups without articulatory suppression than with it, $F(1, 34) = 36.92$, $p < .001$, $\eta^2 = .52$. Finally, there was no interaction between group and condition, $F < 1$.

Concerning the PROMS test, we ran four Bonferroni-adjusted independent samples t-tests (see Figure 2) The t-tests showed that musicians performed better than nonmusicians in all the PROMS subtests (see Figure 2): Melody, $t(34) = 6.09$, $p < .0001$ ($d = 2.03$), Tuning, $t(34) = 4.20$, $p = .0002$ ($d = 1.40$), Speed, $t(34) = 2.50$, $p = .02$ ($d = .83$), and Beat, $t(34) = 3.90$, $p = .0004$ ($d = 1.31$).

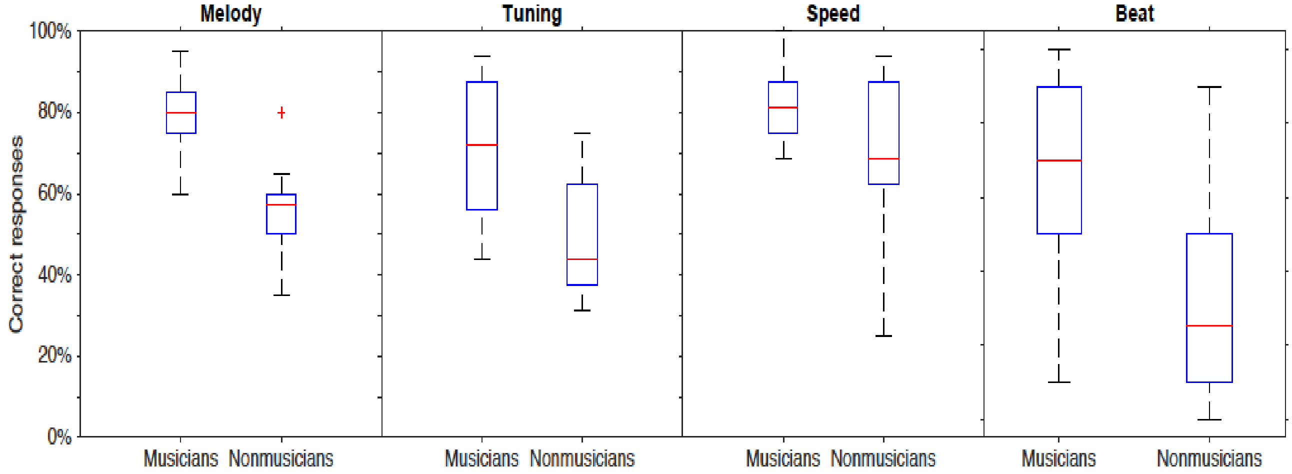

Figure 2. Performance of the PROMS test separately for each subtest. In each box, the red horizontal line indicates the median. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The red crosses are the outliers.

Moreover, we run 24 correlations FDR-adjusted (Benjamini & Hochberg, 1995), with the aim of evaluating if there was a relationship between the four PROMS subtests and the span scores. Table 2 represents these correlations.

Table 2. *Correlations Between the WM Span for Various Modalities and Conditions and the PROMS Subscales*

	Melody	Tuning	Speed	Beat
Span Auditory	.48*	.22	.28	.06
Span Audiovisual	.51*	.35	.39	.22
Span Visual	.16	.34	.17	.01
Span Auditory (S)	.36	.23	.26	.11
Span Audiovisual (S)	.10	.24	.31	.01
Span Visual (S)	.37	.30	.27	.15

Note: Rows labels followed by (S) highlight the span with suppression. *$p < .05$

There were two significant correlations between the subtest Melody of the PROMS test and the span presented auditorily, *r*(36) =.48, *p* = .04, and audiovisually, *r*(36) = .51, *p*=.02, with no articulatory suppression. These correlations showed that the better the performance in the span in these modalities, the better the score in the Melody subtest. There correlation between the melody subtest and the span presented visually was not statistically significant, *r*(36) = .16, *p* > .05. Figure 3 represents these two significant correlations. In addition, we investigated the possible relationship between the digit span and the WAIS-IV performance, but no statistical significantly correlation was found between any of the digit span tests and the WAIS-IV subtests.

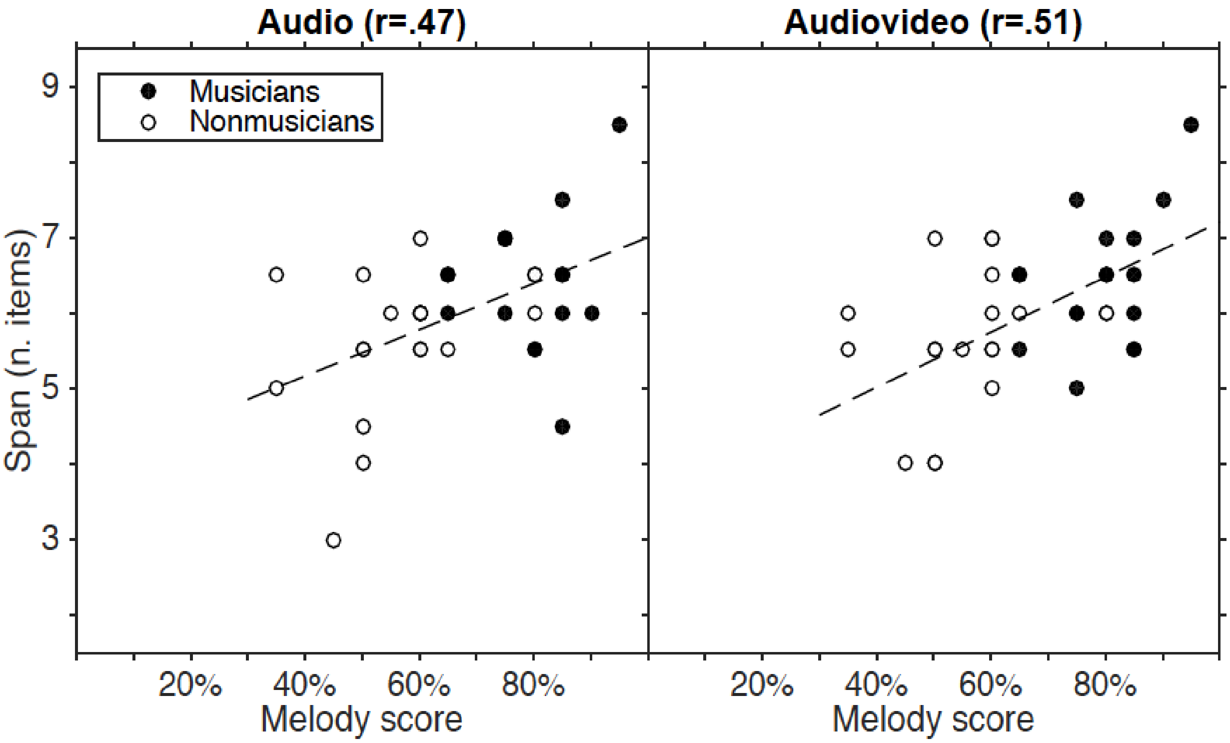

Figure 3. Correlational plots between auditory and audiovisual digit span and the Melody score of the PROMS test. Each dot represents one participant. The graph title includes the value of the correlation coefficient.

Finally, other 12 FDR adjusted correlations were calculated in the musicians' group only, for investigating the relationship between the years of music training received and the span scores, and the hours of weekly practice of the instrument and the span scores. There was a positive correlation of moderate size between the years of music training and the span score in the audiovisual modality with no articulatory suppression, *r*(18) = .48. There was also a moderate negative correlation between the hours of weekly practice of the instrument and the span score in the auditory modality with articulatory suppression, *r*(18) = -.43. Both these correlations became nonsignificant when the p-value was FDR-adjusted.

Discussion

The present study investigated verbal short-term and working memory of musicians and nonmusicians. The three main objectives were: 1) comparing modalities of presentation, by delivering the span auditorily, visually, and audiovisually (e.g., Hansen et al., 2012; Tierney et al., 2008); 2) manipulating the complexity of the task and blocking verbal rehearsal (e.g., Franklin et al., 2008); 3) taking into account the music aptitude of all participants to see whether it could be linked to the memory performance (e.g. Hansen et al., 2012). The digit span was thus delivered in three different modalities, visually, auditorily, and audiovisually. For the second objective, an articulatory suppression condition was added for investigating the role of complexity and of rehearsal strategies. For the third objective, music aptitude was assessed with the PROMS test.

Results revealed that musicians performed better than nonmusicians in the digit span, independently from the modality of presentation and from the condition (i.e., with or without articulatory suppression). This result suggests that there is an advantage of musicians for both maintenance tasks, that tap the passive storage of short-term memory (i.e., digit span forward), and for more complex tasks, which require a higher cognitive load, that tap working memory (i.e., digit span with articulatory suppression). Previous studies showed mixed results, with some that showed a general advantage of musicians in verbal working memory, both in tasks that involved a passive storage and in tasks that involved an active processing of the information to-be-remembered (e.g., Franklin et al., 2008), and with some studies that instead found an advantage in tasks that required only to maintain the information (e.g., Hansen et al., 2012; Lee et al., 2007). The present results support the former findings rather than the latter findings.

The main analysis revealed that musicians performed better than nonmusicians regardless of the modality, even though literature suggested that stimuli presented auditorily are more likely to produce difference between musicians and nonmusicians in memory tasks than visual stimuli. However, only one study before this (Tierney et al., 2008) investigated the different modalities of presentation within the same task-type. Here, although the main ANOVA did not show any modality effect or interaction between group and modality effect, following speculative ANOVAs, run for testing directly the modality hypothesis, suggested that the advantage of musicians over nonmusicians was larger in the auditory and audiovisual modalities with respect to the visual one. Specifically, in the auditory modality, musicians outperform nonmusicians both with and without articulatory suppression, and the distance between groups was large. In the audiovisual modality, instead, the advantage of musician was present only without articulatory suppression. This could be due to the fact that musicians, in the audiovisual presentation without articulatory suppression, could have been focusing mainly on the auditory input, ignoring the visual one (and therefore performing better than nonmusicians, as in the auditory-only condition). Whereas, with the articulatory suppression, they might have decided to focus more on the visual presentation, because of the auditory interference of saying "la la la" aloud while listening to the digits. This could have been an unlucky strategy for them, since their strength relies on the auditory presentation of stimuli, thus dissolving their advantage. The superior performance of musicians in the audiovisual condition (without articulatory suppression) might have another explanation. They could be, for example, better in integrating information coming from different sensory modalities, since the multisensory integration is something that musicians do while playing the instrument (e.g., see Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012; Paraskevopoulos, Kraneburg, Herholz, Bamidis, & Pantev, 2015). To sum up, even though our second analysis did suggest a superiority of musicians specifically in auditory and audiovisual modality, the modality hypothesis was weakly supported.

Another aim of the present study was to look at the possible relationship between specific music skills and working memory. Correlations between auditory and audiovisual span and the melody subtest of PROMS were found. One possible explanation of these correlations concerns the nature of the melody task. This task requires to listen to a melody, presented twice, keep it in mind, and then judge whether a third melody is the identical or not to the melody presented twice. One of the parameters for increasing the difficulty of this subtest is increasing the number of notes (i.e., as for the numbers in the digit span). Therefore, memory is involved. This explanation, though, is not sufficient. In fact, there was no correlation with the visual digit span, so if memory was responsible of these correlations, we should have found also a correlation with the visual span. Therefore, this result suggests that the modality of presentation of verbal stimuli (i.e., visual and auditory) should be carefully considered, because they can rely on different memory subsystems. The present study was not the only one to observe a correlation between span and a specific music aptitude skill. In fact, Hansen and colleagues (2012) found a correlation between the forward digit span and one subtest of the MET (Wallentin et al., 2010), the rhythm subtest. In this subtest, difficulty was manipulated by increasing the number of tones composing each rhythmic pattern (i.e., from 4 to 11). Nevertheless, also the MET had a melody subtest, but there, no correlation with the span was found (Hansen et al., 2012).

Finally, one possible explanation of the superiority of musicians over nonmusicians in the digit span, could be that musicians have better and more performing general cognitive abilities than nonmusicians (e.g., Schellenberg, 2004). In the present study, we used two subtests of the WAIS-IV (i.e., Visual Puzzles and Vocabulary) as control measure for general cognitive abilities. No significant difference between musicians and nonmusicians was observed in these two subtests.

To sum up, the current results support an advantage of musicians over nonmusicians in verbal short-term memory and working memory, independently from the articulatory suppression condition (i.e., use of better rehearsal strategies could not be the explanation of the advantage). Modality seems also to play a role, but this role looks small. Furthermore, the correlation between the melody subtest and the digit span in auditory and audiovisual modality suggest that some music skills might be related to the ability of memorizing sequences of digits, and it suggests that this relationship is modality-dependent.

Study 3. Can you read the score? Short-term and Working Memory of Trained Musicians, Self-taught Musicians, and Nonmusicians.

Aims

STUDY 1 and 2 evidenced an advantage of musicians over nonmusicians, especially in short-term memory and working memory tasks with tonal and verbal material. We were particularly interested in the advantage in the verbal domain, because it is a component that is not directly trained in the music training. In the present study we thus aimed to investigate again the verbal working memory of musicians, in comparison with nonmusicians, but introducing a new group of musicians: musicians who cannot read music notation. The idea was to understand better which part of the music training is linked to an enhanced memory performance: is it being a musician in general? Is it learning to read the music notation and to play while reading a score? Or is it being able to play an instrument in general, even without being able to read music notation? The hypothesis was that learning to translate a symbol (i.e., the note) into a sound, through a specific motor response, could be the key of the relationship between music training and enhanced verbal memory. Learning to play an instrument (excluding the motor aspect) is like learning to read a foreign language. As the person learns a new language, s/he could also learn specific strategies to remember, for example, the association between a symbol and a sound. The same could happen in music, where the musicians could learn specific strategies to remember the correspondence between music notation, sounds, and motor actions. Therefore, we hypothesize that training in reading music notation can help developing specific strategies that can be useful in the learning of a language too, and consequently, this could be linked to also verbal memory performance. For this reason, a group of musicians that could not read music notation was recruited for the study together with a group of musicians that could read the music score. Groups were recruited by matching as best as possible the level of training, therefore the everyday practice (hours per week played) and years of music experience. A control group of nonmusicians was also included. General cognitive abilities were controlled for with two subtests of the WAIS-IV. The music aptitude of all participants was assessed with the PROMS test. The task used to tap memory was, alike STUDY 2, the digit span, but unlike STUDY 2, we presented the digit span in a forward and in a backward condition, to have both a measure of the capacity of short-term memory (forward span) and a measure of working memory (backward span). For explorative purposes, two final questionnaires were also administered to investigate the mental imagery ability of all participants, specifically for visual mental imagery (e.g., imaging a sunset) and auditory mental imagery (e.g., imaging the sound of church bells). The aim of these questionnaires was to investigate whether the performance of the span was correlated also to the ability of creating mental images (that could be useful in using mental strategies during the memory task).

Method

Participants

One hundred and two young adults participated to the study. The mean age was 22,7 years (min = 19; max = 33). There were 35 females and 65 males. Participants were: (1) thirty-six conservatory students, music schools' students, and/or professional musicians; (2) thirty-three self-taught musicians that could not read music notation (apart from the basics learnt at school); (3) thirty-three nonmusicians, that never underwent a music training, or took minimal years of music lessons when they were children (i.e., less than 2 years). From now on, musicians that received a formal training will be called "readers" and self-taught musicians will be called "nonreaders", to emphasize the main difference between the musicians' groups, which was being able to read or not a music score. Demographic details are reported separately for each group in Table 1. Reader musicians received slightly more years of training with respect to nonreader musicians, and they played substantially more hours per week with respect to the nonreaders. All participants had normal audiometric thresholds (assessed with an audiometry, for frequencies of 500, 1500, and 4000 Hz) and reported normal (or corrected to normal) vision.

Table 1. *Age, Education, Performance (Raw Scores) in the WAIS-IV Visual Puzzles and Vocabulary Subtests, music education and practice hours.*

	Readers *N* =36 (13 females)	Nonreaders *N* = 33 (4 females)	Nonmusicians *N* = 33 (18 females)	
Age (yrs)	22,7 (2,1)	23,2 (2,1)	22,1 (1,6)	*p* = .137
Education (yrs)	16,3 (1,9)	16,5 (1,6)	15,9 (1,4)	*p* = .333
Visual Puzzles (max score 26)	17,53 (4,34)	17,48 (3,72)	15,42 (3,1)	***p* = .037**
Vocabulary (max score 57)	45,75 (8,14)	47,18 (5,97)	44,76 (6,03)	*p* = .354
Music Training (yrs)	11,8 (3,7)	10,2 (2,3)		***p* = .032**
Weekly Practice (hrs)	14,2 (9,2)	7,2 (4,9)		***p* < .001**

Note. Mean (SD). In the last column on the right, the p-value reflects to the analysis between (t-test) or among groups (ANOVA). In bold characters, significant p-values. In the Visual puzzles, post-hoc comparisons revealed that the groups that differ significantly were readers > nonmusicians, *p* = .023, and nonreaders > nonmusicians, *p* = .018.

Materials

Materials were largely similar to the STUDY 2, with a few differences described below.

Digit span test. Digits were presented in audio, video and audiovisual modality, alike in STUDY 2. The pool of digits presented comprised digits from 1 to 9, starting with a 3 digits sequence in the forward condition, and with 2 digits in the backward condition. Sequences were fixed for all participants and were taken from the WAIS-IV Digit Span subtest, and here adapted to be presented with a computer. The digit span was administered in two different conditions, forward and backward. The backward condition here replaced the articulatory suppression used in STUDY 2. In the forward condition the participant had to type the sequence of digits previously saw/heard in the same order of presentation. In the backward condition the participant had to type the sequence starting from the last digit and going backward to the first one presented.

Alike in STUDY 2 the Mini PROMS was test used to assess the music aptitude of all participants, and two control measures were administered to tap general cognitive abilities WAIS-IV. The Visual Puzzles subtest was used to control for the nonverbal reasoning of participants, the WAIS-IV. The Vocabulary subtest was used to control for the verbal *semantic* skills of participants. Finally, two new, exploratory questionnaires were introduced in the study: the mental imagery questionnaires. These self-report questionnaires were used to investigate the mental imagery ability of participants. I used both the visual mental imagery questionnaire (the VVIQ, Vividness of Visual Imagery Questionnaire, Marks,

1973) and the auditory mental imagery questionnaire (build ad hoc basing on the structure of the VVIQ). Each questionnaire was composed by 16 items, divided into four sets of questions. Each set of questions was referring to a specific object that had to be imagined, and each individual item of the set asked to imagine a particular detail of the object. Participants had to imagine the scenes first with open eyes, and successively with closed eyes. Next to each item, there were two empty boxes (one for the open eyes condition, and one for the closed eyes one) in which the participant had to fill in a number from 1 to 5, depending on how much the mental image created was vivid (1 “not able to create the image”, 5 “image definitely vivid, like it is really seen/heard”).

Apparatus & procedure

The apparatus used was identical to STUDY 2. The procedure was identical to STUDY 2. Half of subjects (of each group) started with the forward condition, whereas the other half started with the backward condition. The mental imagery questionnaires were administered at the end of the experiment.

Analyses

The analyses will be divided into two sections, the main analysis, which was hypothesis driven, and the exploratory analysis. In the main analysis the possible effect of the group (i.e., reader musicians, nonreader musicians, and nonmusicians) on the digit span was investigated through a repeated measure ANOVA, separately for the forward and the backward span. The ANOVAs here performed have the modality (i.e., auditory, audiovisual, visual) as within factor, and the control measures in which the groups performed differently (i.e., Visual Puzzles) as covariate. In the main analysis, an ANOVA on the PROMS scores was also run. Correlations among these measures were also performed.

The exploratory analysis focused on the mental imagery questionnaires, through a repeated measures ANOVA that included the group as between factor, and the modality (i.e., visual imagery, auditory imagery) and the condition (i.e., open eyes, closed eyes) as within factors. Correlations were run, to see whether the ability of creating vivid images is related to the performance in the digit span tests.

I will report all the statistics for the main effects and the statistics for only the significant interactions; only significant correlations will be reported for sake of brevity. Means and standard deviations will be reported only when there will be not any graphical representation. For multiple comparisons, I will report p-values both without and with FDR correction.

Results

Main analyses

Digit span Forward. A first repeated measures ANOVA with the span forward score (i.e., number of correct sequences reproduced) as dependent variable, and with the modality (i.e., auditory, visual, and audiovisual), and the group (i.e., readers, nonreaders, and nonmusicians) as independent variables was run. There was not a significant effect of group, $F(2,99) = .25$, $p = .78$, $\eta^2 = .005$, meaning that overall there was no difference among the three groups in the digit span performance. There was a significant effect of modality, $F(2,198) = 4.58$, $p = .011$, $\eta^2 = .04$, specifically, the performance of all subjects was slightly better with digits presented auditorily ($M = 10.78$, $SD = 2.33$) than visually ($M = 10.11$, $SD = 2.58$), t (200) = -1.96, $p = .051$, but no difference between the audiovisual presentation ($M = 10.74$, $SD = 2.43$) and the visual presentation, $t(201) = -1.82$, $p = .071$, and between audiovisual and auditory modality, $t(202) = -.12$, $p = 906$, was found. When the FDR correction was applied, the difference between auditory and visual modality was no longer significant ($p = .106$). A significant interaction between modality and group was observed, $F(4, 198) = 3.83$, $p = .005$, $\eta^2 = .07$, suggesting that the three groups performed differently in the various modalities. Post--hoc t-tests revealed that reader musicians outperformed nonmusicians only in the audiovisual modality, $t(67) = 2.01$, $p = .048$, $d = .49$ (g = .48), but the difference between groups was anyway little (i.e., on average one correct response). There were no others significant effects. Here again, when correcting the p-value for multiple comparisons, the advantage of readers musicians in the audiovisual span became no longer significant ($p = .433$). Figure 1 represents the span performance in the forward digit span.

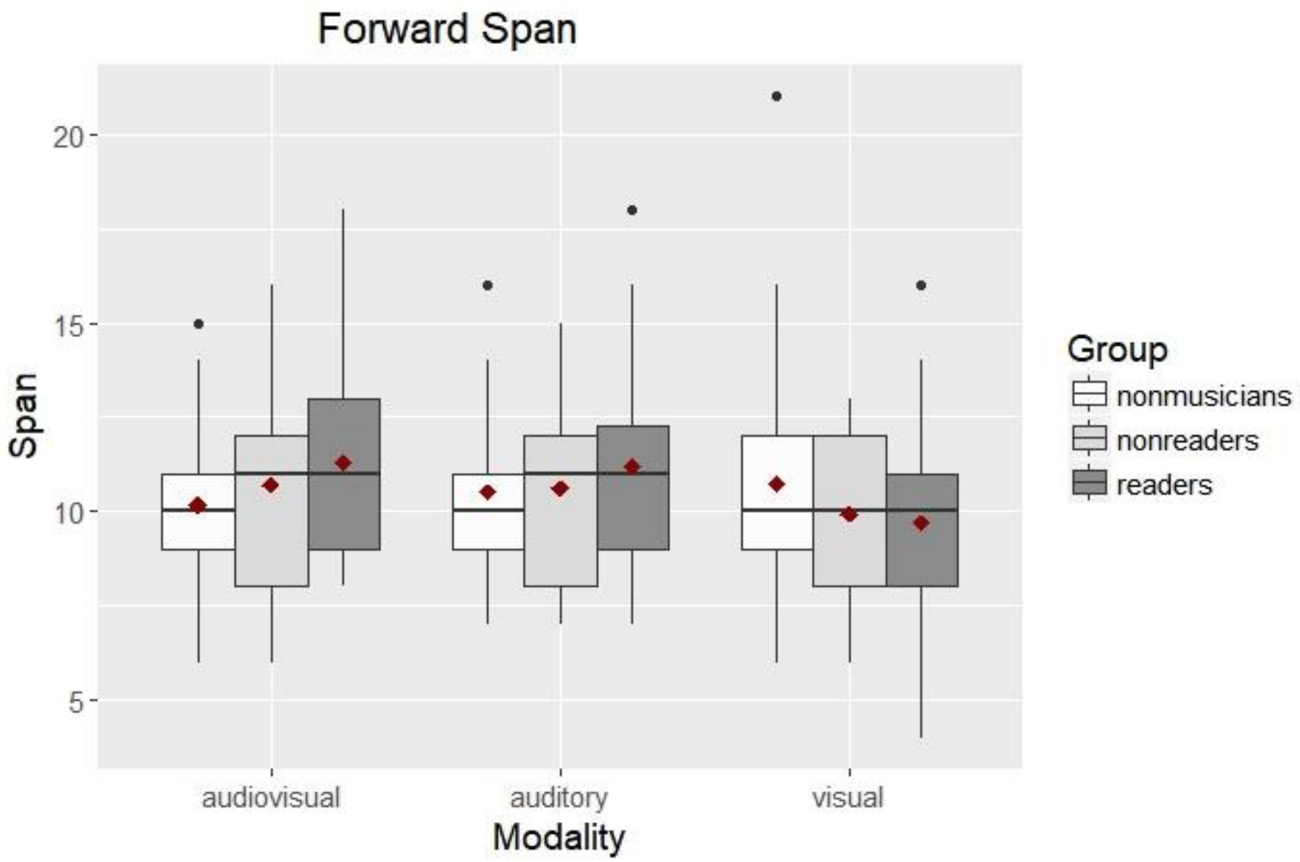

Figure 1. Digit Span Forward performance (i.e., number of sequences correctly reproduced) separately for modality of presentation and group. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The black dots are outliers[1].

---

[1] The outliers in the different modalities, within the same group, are different subjects (e.g., it was not a unique subject that had a higher performance in several conditions)

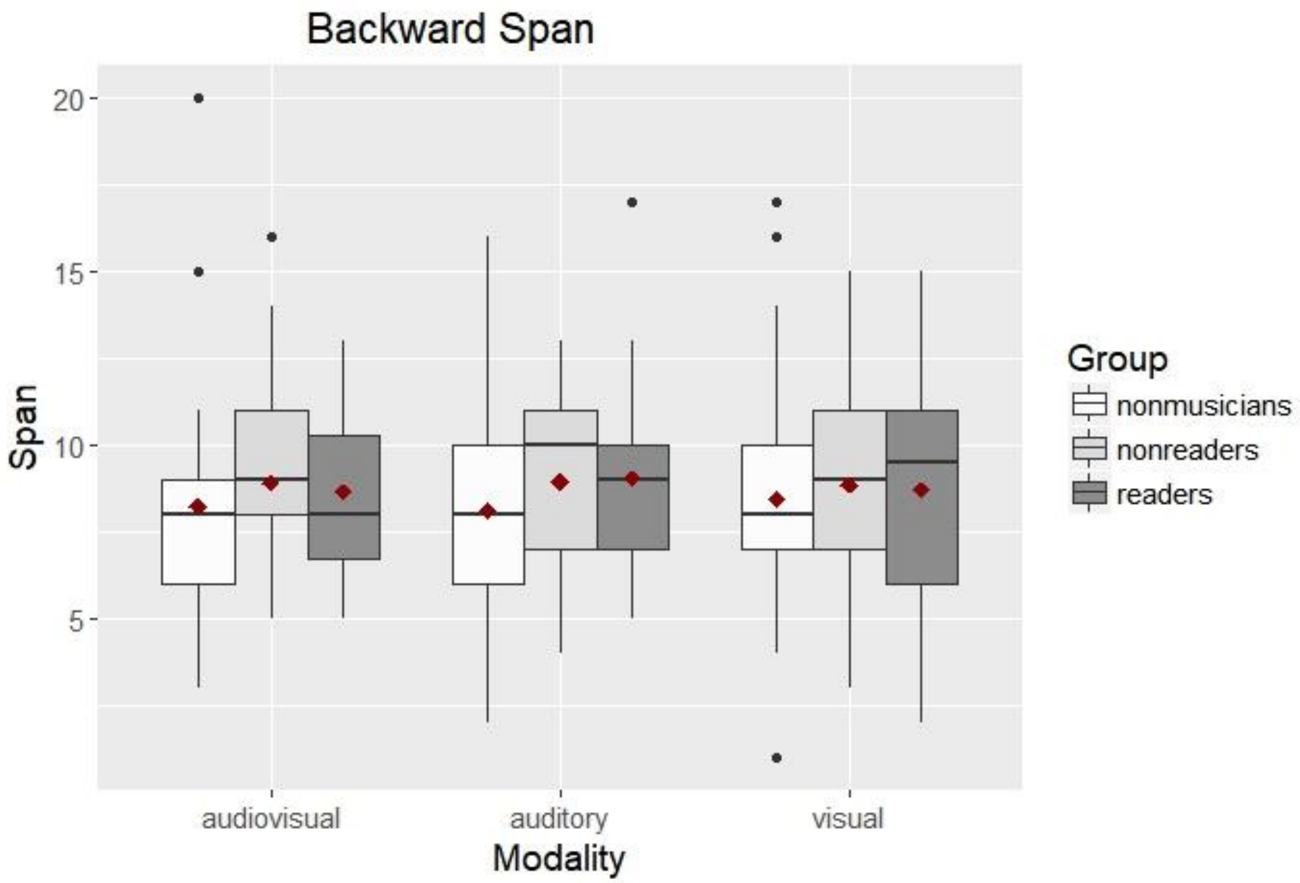

Figure 2. Digit Span Backward performance (i.e., number of sequences correctly reproduced) separately for modality of presentation and group. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The black dots are outliers.

Because the analysis for the control measures (i.e., the WAIS-IV subscales) revealed that readers and nonreader musicians outperformed nonmusicians in the Visual Puzzles subtest, this measure was included as a covariate and a further ANOVA was calculated. The covariate was not significant, *F*(1, 98) = *.002, p = .963,* η^2 *= .001*, and after its inclusion in the analysis, the interaction between group and modality was still significant *F*(4, 196) = 3.56, *p* = .008, η^2 = .07, suggesting that the small difference among groups observed in some modalities, it cannot be explained by a general superiority in the nonverbal reasoning test.

Digit span backward. A repeated measures ANOVA with the span backward score (i.e., number of correct sequences reproduced) as dependent variable, and with the modality (i.e., auditory, visual, and audiovisual), and the group (i.e., readers, nonreaders, and nonmusicians) as independent variables was run. The Maulchy test revealed that the sphericity was violated, therefore a Huynh-Feldt correction was applied. The group was again not significant, $F(2,99) = .59$, $p = .555$, $\eta^2 = .01$. The modality, here, was not significant, $F(1.9,189.5) = .09$, $p = .906$, $\eta^2 = .001$. Differently from the forward span, here there was not any interaction between modality and group, $F(3.8, 189.5) = .395$, $p = .804$, $\eta^2 = .01$. Figure 2 represents the span performance in the backward digit span. When introducing the Visual Puzzles as covariate, this was significant, $F(1,98) = 5.02$, $p = .027$, $\eta^2 = .05$, suggesting that the performance in this test is related (to a small degree) to the performance in the digit backward span. No other interactions were significant.

To further explore the relationship between the score in the visual puzzles and the digit span performance, some correlations were run. There were two small correlations, with the backward auditory span, $r(102) = .25$, $p = .01$, and with the backward audiovisual span, $r(102) = .28$, $p = .004$, and no correlations between the digit span forward and this WAIS-IV subtest were found. Moreover, the readers group and the nonreaders group differed significantly in term of years of training and hours of weekly practice; therefore, the span performance of reader musicians and nonreader musicians was correlated to these two variables to check whether they could be linked to the memory performance, but no correlations were found.

PROMS test. A one-way ANOVA with the total score of PROMS test as dependent variable and the group as between factor[2] revealed that the group was significant, $F(2, 99) = 32.17$, $p < .001$, $\eta^2 = .39$. Namely, readers musicians and nonreader musicians performed better in this test than nonmusicians (see figure 3). When the Visual Puzzle score was included as a covariate, it resulted significant, $F(1, 98) = 5.63$, $p = .02$, $\eta^2 = .04$, but the group factor remained significant, $F(2, 98) = 27.17$, $p < .001$, $\eta^2 = .34$. Post-hoc tests revealed that reader musicians outperformed nonmusicians, $t(67) = 7.51$, $p < .001$, and that nonreader musicians outperformed nonmusicians too, $t(64) = 5.96$, $p < .001$. A significant difference between readers and nonreaders was also found, with the former performing just slightly better than the latter, $t(66) = 1.987$, $p = .05$. When the p-values were FDR corrected, the significance did not change ($p \leq .05$). Moreover, the total score of PROMS was positively correlated with the Visual Puzzles test, $r(102) = .33$, $p < .001$, as expected from the significant covariate in the ANOVA.

---

[2] Other four one-way ANOVAs for each one of the PROMS subtests were run, and in each one of them the group factor was always significant ($p < .05$)

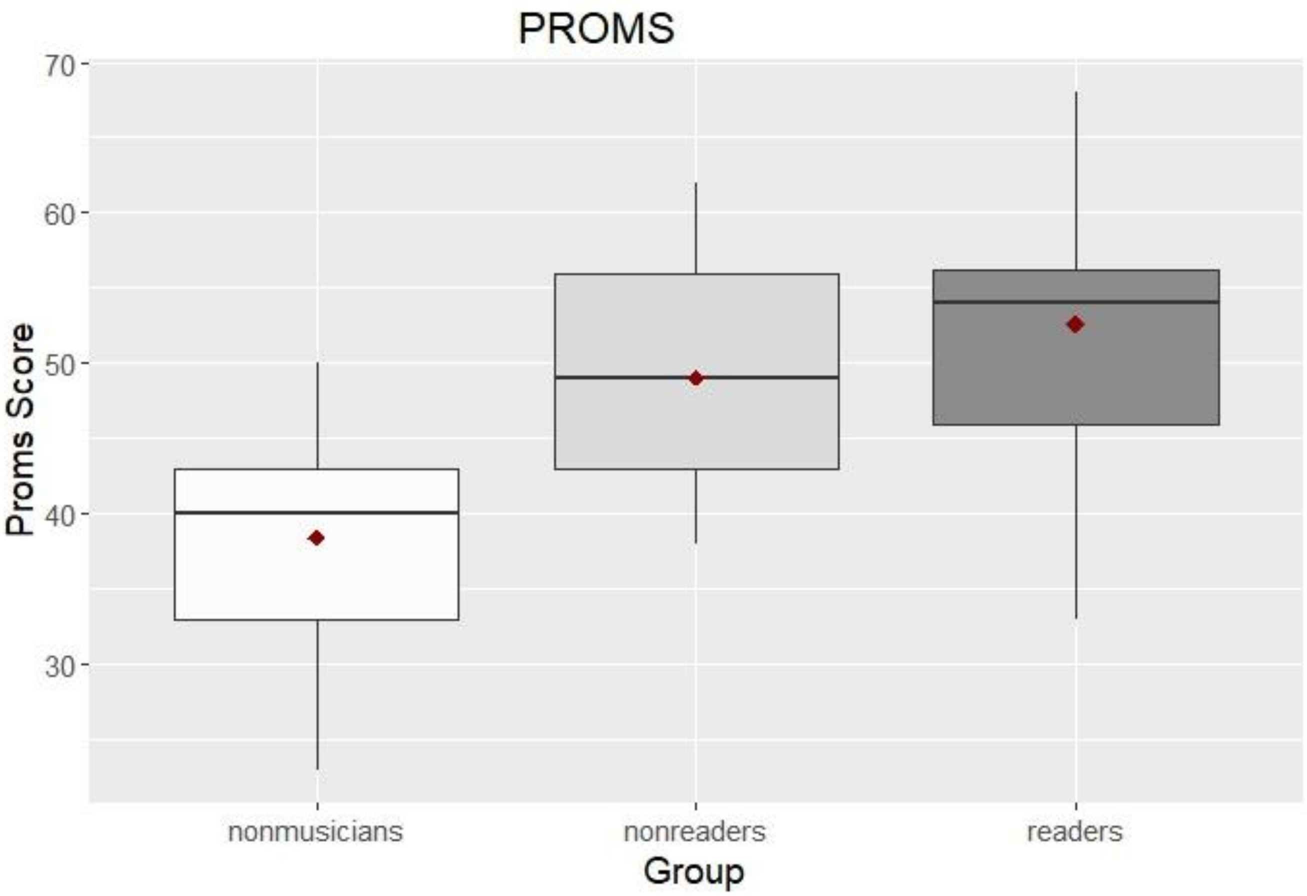

Figure 3. PROMS test total score, separately for nonmusicians, reader and nonreader musicians. Inside each box, the horizontal line indicates the median and the red diamond the mean. The edges of the box's edges represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%.

Since the readers group was the one with highest score in the PROMS test, and since they had also more years of music training and much more hours of weekly practice than nonreaders, some correlations between years of music training, hours of weekly practice and PROMS were run. Correlations are reported in Table 2.

Table 2. *Pearson correlations between years of music training, hours of weekly practice, and PROMS test*

	MELODY	TUNING	SPEED	BEAT	TOT PROMS
Years of music training	0.447 ***	0.637 ***	0.371 ***	0.380 ***	0.611 ***
Hours of weekly practice	0.322 ***	0.328 ***	0.123	0.213 *	0.335 ***

Note. *p <.05, **p < .01, ***p< .001

Correlation between the PROMS test (both the total score and each subtest) and the Span score were also run. A few positive correlations were found: alike STUDY 2, a positive correlation between the melody subtest and the forward auditory span, *r*(102) = .19, *p* = .05, and between the melody subtest and the forward audiovisual span, *r*(102) = .20, *p* = .041. Both correlations, however, were weaker than those observed in STUDY 2. Other further positive correlations between the beat subtest and the forward audiovisual span, *r*(102) = .22, *p* = .029, and the beat subtest and the backward auditory span were found, *r*(102) = .22, *p* = .029. Finally, a negative correlation between the tuning subtest and the forward visual span emerged, *r*(102) = -.21, *p* = .038.

Exploratory analyses

Imagery questionnaires. An ANOVA with the score at the imagery questionnaires as dependent variable and condition (i.e., eyes opened, eyes closed), modality (i.e., visual imagery, auditory imagery), and group as independent variables was calculated. There was a significant effect of the condition, *F*(1, 99) = 43.23, *p* <.001, η^2 = .29, meaning that the vividness of the images created was higher when the task was performed with closed eyes (*M* = 123.5, *SD* = 18.59) than with open eyes (*M* = 114.5, *SD* = 18.26). There was a significant interaction between modality and group, *F*(2, 99) = 4.2, *p* = .018, η^2 = .08, meaning that the three groups performed differently depending on the imagery modality of the questionnaire (see Figure 4). When adding the Visual Puzzles as a covariate, the covariate resulted nonsignificant, *F*(1, 98) = .50, *p* = .481, η^2 = .005, and the interaction remained significant, meaning that this result did not depend on the difference among groups in nonverbal reasoning abilities, *F*(2, 98) = 3.59, *p* = .031, η^2= .07. Post hoc analysis revealed that readers differed significantly from the nonmusicians group only in the auditory imagery score, *t*(67) = 2.54, *p* =.013, and not in the visual one, *t*(67) = 1.28, *p* = .206. As the readers, nonreaders differed significantly from nonmusicians in the auditory imagery questionnaire, *t*(64) = 3.55, *p* <.001, but they did not differ in the visual imagery questionnaire, *t*(59) = 1.58, *p* =.120. Finally, readers and nonreader musicians did not differ in the auditory questionnaire, *t*(67) = -.9, *p* = .371, nor in the visual questionnaire, *t*(64) = -.42, *p* = .679. The difference between readers and nonmusicians in the auditory questionnaire remained significant after correcting for FDR the p-value (*p* = .039), as well as between nonreaders and nonmusicians (*p* = .004). This suggests that musicians, especially nonreaders, could create more vivid auditory mental imagery than nonmusicians.

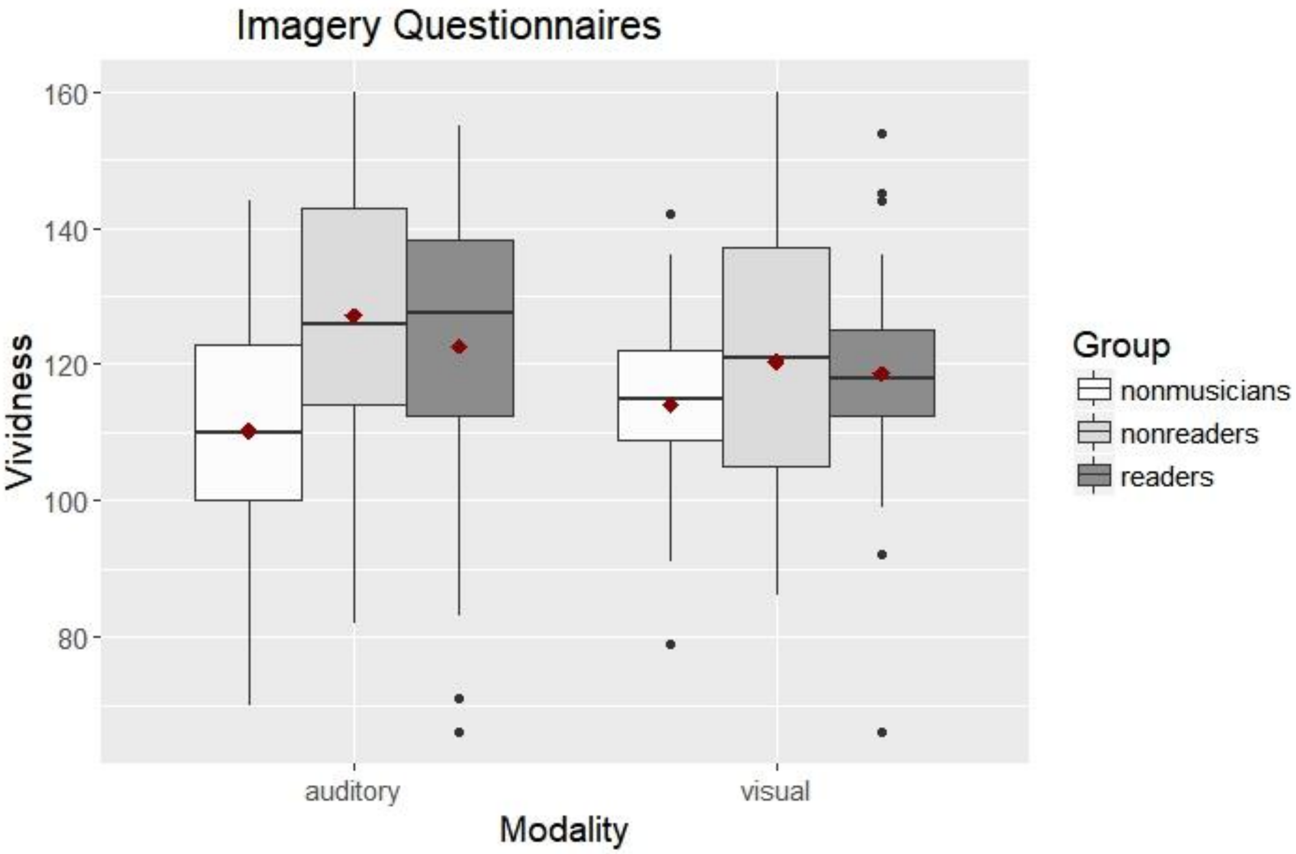

Figure 4. Self-report scores in the mental imagery questionnaires: the higher the value, the higher the vividness reported. For each modality, the score for the opened eyes condition and closed eyes condition was combined, to evidence the interaction between group and modality. The maximum score obtainable is 160. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The black dots are outliers.

Finally, to see whether imagery skills could be linked to the performance in the digit span, some correlations between the imagery questionnaires and the digit span score were run. No significant correlations were found, indicating that being able to create more vivid mental images does not help the performance in the memory task.

Discussion

The present study aimed to investigate the verbal short-term and working memory skills of musicians and nonmusicians focusing on different groups of musicians: reader musicians, that had a formal music training, and nonreader musicians, that were self-taught and who reported being not able to read music notation. A forward and a backward digit span were administered in three modalities, auditorily, visually, and audiovisually, to take into account a possible selective advantage of musicians with auditory digits, as suggested by STUDY 2. Music aptitude was also assessed, with the PROMS test, to have an objective measure of music perception skills of all participants and see whether they could be related to the span performance. For exploratory purposes, auditory and visual imagery skills were also investigated through two self-report questionnaires, with the aim of seeing if the performance in a working memory test could be related to high imagery skills. In other words, explore whether a person who can have vivid mental images can exploit this ability when remembering a sequence of digits.

The results of the main analysis in the forward digit span revealed a general (small) effect of the modality. The performance of all participants was slightly better when digits were presented auditorily than visually, but this difference was no longer significant when correcting the p-value for multiple comparisons. Concerning the group, we expected to find a difference between reader musicians and nonmusicians particularly in the forward digit span, since the meta-analysis (STUDY 1) suggested that there is a moderate advantage of musicians over nonmusicians in this task. We did not find a general effect of group, but an interaction between group and modality emerged in the span forward. This meant that there were differences among groups depending on the modality of presentation of the digits. Post-hoc tests revealed that reader musicians outperformed nonmusicians only in the audiovisual presentation of the digits. This difference was however small and when looking at the total number of correct responses, reader musicians made on average one more correct response than nonmusicians. This was the only significant difference among groups observed, however, qualitatively we could observe some other differences among groups also in the auditory modality (See Figure 1 and Figure 2), but not in the visual one, where all the three groups performed similarly, or even nonmusicians seemed to perform better than the other groups. However, the fact that with the p-value correction for multiple comparisons the difference between reader musicians and nonmusicians was no longer significant, supports the fact that these differences are rather small. Note that here the effect size, expressed with the d of Cohen, was .47, that is, close to be moderate. This suggest that even if the effect sizes included in the meta-analysis are moderate, in terms of distance between groups in the raw score of the memory tasks, we are talking about small numbers (e.g., on average, 1 or 2 correct responses more).

Concerning the backward digit span, here, the only significant effect was that of the covariate. Even if readers and nonreaders had higher scores than nonmusicians in the Visual Puzzles subtest, the group was not significant, but qualitatively a small difference among groups (with readers and nonreaders performing slightly better) was observed. Moreover, two small correlations between the Visual Puzzles and the backward auditory and audiovisual span were found. This is in line with the literature that suggests that a relationship between working memory and general cognitive abilities, such as fluid intelligence, exists (e.g., Ackerman, Beier, & Boyle, 2005; Giofrè, Mammarella, & Cornoldi, 2013). The fact that the group was not significant, however, is against the evidence of STUDY 1, which reported a moderate difference between groups in verbal working memory tasks. It is worth noting that, even if the meta-analysis supported this advantage, there was higher heterogeneity in the verbal working memory tasks results in comparison to short-term memory tasks, where no heterogeneity was observed. This means that when investigating working memory in musicians and nonmusicians, results tend to vary. This could depend also on the fact that the tasks that investigate working memory are various, and they might tap slightly different aspects. A study that found similar results was the one by Hansen and colleagues (2013), in which the authors observed an advantage of musicians over nonmusicians in the forward digit span but not in the backward span.

The present results are interesting for several reasons: first of all, they are not completely in line with the results of STUDY 2, that used a very similar paradigm. In STUDY 2 a general advantage of musicians was found, and explorative analysis suggested that the advantage was present mainly in the auditory and audiovisual modality but there was no interaction between group and modality. The difference between the present results and STUDY 2 results could be due to several reasons, for example the number of participants, which here is larger, and the number of groups, which is three instead of two. In any case, it is difficult to tell whether the different results between the present study and STUDY 2 are due to these differences. Secondly, here a difference between musicians and nonmusicians in the Visual Puzzles subtest was found, unlike STUDY 2. The fact that here musicians, both readers and nonreaders, showed better nonverbal reasoning skills than nonmusicians (i.e., higher score in the Visual Puzzles), is not always supported by literature. As we saw in the third chapter, some studies found that musicians have higher IQ scores than nonmusicians (e.g., Schellenberg, 2011; Schellenberg & Mankarious, 2012; Gibson, Folley, & Park, 2009), and even if IQ measures include more than one task, nonverbal reasoning is often one of them. Here, anyway, the Visual Puzzles performance was not linked to the small advantage of musicians over nonmusicians observed, and this is in line with other studies on working memory which found specific advantage of musicians over nonmusicians even when intelligence was held constant (Jakobson et al., 2008; Roden et al., 2014), or even when the control measures for general cognitive abilities did not evidence any difference between groups (Franklin et al., 2008). Concerning the backward digit span, here the only significant factor emerged was the covariate (i.e., the Visual Puzzles); this result is in line with the literature, which suggests that there is a relationship between working memory and general cognitive abilities, such as fluid intelligence (e.g., Ackerman, Beier, & Boyle, 2005; Giofrè, Mammarella, & Cornoldi, 2013). Moreover, two small correlations between the Visual Puzzles and the backward auditory and audiovisual span were also found.

Concerning the difference between readers and nonreaders, the present study was the first to choose these particular groups of musicians, therefore our hypotheses were speculative and not based on previous results. Note that there are previous studies which compared cognitive skills of different classes of musicians , but usually the variable manipulated was the instrument played (e.g., Pellicano et al., 2010) or the distinction between experts and amateurs (e.g., Hansen et al., 2013; Travis, Harung, & Lagrosen, 2011). The choice of having these two particular groups was made because, if the musicians who had a formal training (i.e., readers) would show an advantage over musicians who are self-taught and cannot read music notation, this could reveal that the advantage of musicians observed in literature in these tasks belongs only to individuals who underwent a formal training, specifically, who learnt to read music notation. However, the present study did not reveal any difference between the two groups, suggesting that being able of reading the music notation is not a key factor for performing better in memory tasks. Moreover, the ideal sample of self-taught musicians should have had the same amount of years of music training and should have played a comparable number of hours per week. This was not the present case, in particular for the large difference emerged in the everyday practice of the two groups of musicians. Due to this discrepancy, the possible influence of years of music training and hours of weekly practice was investigated, by running some correlations with the digit span but no correlation was found. This lack of correlation suggests that these variables were not related to the performance in the digit span, therefore we should not expect a difference depending on the years of training received. The effect of the number of years of training undergone could be better assessed by including also people who had just began the music training; otherwise, having all musicians that are experts (around 10 years of music training), could make it difficult to detect any difference linked to the amount of training undergone. To sum up, the present results seem to reflect a general advantage of musicians in auditory working memory tasks, as emerged by the interaction between group and modality in the first ANOVA; but this advantage, however, is small, as also reflected by the effect-size. Moreover, the t-tests here ran as post-hoc analyses showed only a small advantage of reader musicians in the forward audiovisual span, but this difference was no longer significant when correction for multiple comparison was applied, reinforcing the idea that, even if an interaction emerged, this was a weak effect.

Concerning the PROMS test, this measure revealed to be a reliable test to distinguish between musicians and nonmusicians, as observed in STUDY 2, and here it distinguished also between readers and nonreaders. When looking at a relationship between music aptitude score and digit span performance through a correlation analysis, alike STUDY 2, two small correlations between the melody subtest and the auditory and audiovisual forward span emerged. This, as already discussed in STUDY 2, could be due to the fact that the melody subtest requires to memorize a melody (which is made by a sequence of tones) and that, alike in the digit span, the difficulty of the test is manipulated by increasing the number of notes in the melody. But again, the fact that the correlation was not found with the visual span suggests a separate system for auditory memory (both for verbal and tonal stimuli) and visual memory. Moreover, here two further correlations between the PROMS beat subtest and both the forward auditory span and the backward audiovisual span were found, alike the study by Hansen and colleagues (2012), in which the authors observed a correlation between a subtest of a music aptitude test that tapped the rhythmic skills of participant and the span forward (which was presented only auditorily). To sum up, it is very much likely that some parts of music aptitude tests involve memory, and that correlations emerge with the subtest that are more similar to the memory task presented (e.g., the melody subtest is composed by various number of tones, the digit span is composed by various number of digits: in both tests, each element has a specific identity in the sequence).

The exploratory analysis showed another interesting result: that musicians, both readers and nonreaders, reported to create more vivid mental images than nonmusicians, but only when these images were auditory. Nevertheless, this ability seems to be not responsible of their better performance in the auditory span, because no correlations between the mental imagery questionnaires and span scores were found. The idea that the ability of creating mental vivid images could help the use of mental strategies in the span, and therefore improve the performance, it is not supported by the present results.

To conclude, the present results showed no difference between musicians who received a formal music training (i.e., readers) and self-taught musicians (i.e., nonreaders) in the memory task. However, reader musicians performed slightly better than nonmusicians in the digit span forward, but only when the digits were presented audiovisually. This advantage is rather small, (about one more correct answer on average), and it seems to be not explained by a general difference in nonverbal reasoning abilities. Moreover, the difference became nonsignificant when correcting the p-value for multiple comparisons. No difference between groups was found in the backward digit span, which requires a greater cognitive load and a manipulation of the information. The fact that here results were weak and different from what emerged in STUDY 1 and 2 might be caused by the many variables here included, which could have produced a low statistical power. Nevertheless, the present study had a large sample (i.e., 102) in comparison to most of other studies present in literature, and it could be also possible that the significant effects are reduced for this reason (Camerer et al., 2018). Here the aim was to compare two different classes of musicians, with the substantial difference of being able to read or not music notation. We failed with the goal of having groups with similar expertise (i.e., there was a substantial difference in the hours of weekly practice). This limitation stresses the difficulty of running quasi-experiments trying to control possible confounding variables. However, the fact that the only difference found in the post-hoc tests was between the reader musicians and the nonmusicians suggest a formal training (and therefore being able to read music notation) could be important in determining an advantage of the musicians. A further limitation of the present study is that, again, it is not possible to attribute the differences that emerged among groups to the music training, or to other variables, such as being involved in an activity that stimulates several senses and requires attention and commitment, and/or personality traits, such as creativity, that are often present in people who are devoted to art.

Study 4 - Performance of Musicians and Nonmusicians in a Short-Term Memory Task with Different Materials

*Study conducted at the Centre de Neuroscience de Lyon (CRNL) under the supervision of Barbara Tillmann and Anne Caclin

Aims

STUDY 1 (i.e., the meta-analysis) showed an advantage of musicians mainly in musical and verbal memory tasks. The present study aimed to investigate further these advantages. First of all, we wanted to understand whether the musicians' advantage was linked both to auditory stimuli and visual stimuli. For example, verbal stimuli are often presented auditorily, and so it is difficult to determine whether this advantage is driven from a superior auditory short-term memory, or if it is linked to verbal stimuli per se. We wanted to further investigate this by presenting verbal stimuli visually and see whether there is still an advantage of musicians over nonmusicians. Note that this was already done in Study 2 and in Study 3, and results suggested that the advantage of musicians was more pronounced in the auditory and audio-visual presentation of the stimuli (i.e., sequences of numbers), and not as much in the visual one. Here, we chose to use syllables instead of numbers because some studies suggest that numbers are processed differently from other verbal stimuli (e.g., syllables, words, nonwords). For example, children with mathematical difficulties have a selective impairment in short-term memory for numbers, but not for words (see Passolunghi & Siegel, 2001)

Furthermore, we wanted to see whether the superiority of musicians extends to other kind of stimuli that are not verbal (e.g., words, syllables, numbers) and not musical (e.g., short melodies), both auditory and visual, because, to our knowledge, no previous study has done that. In particular, for the nonverbal material, we wanted to have two conditions: a contour one, and a no-contour one. In music, the pitch contour is the up and down pitch variation created by a melody but regardless the exact pitch-value of each individual note. In other words, the pitch contour is the "higher/lower" variation that is perceived along the pitch continuum when we listen to music. Here, "contour" stimuli were defined as sequences of elements that are related to one another along a physical continuum, elements that are in a relationship with the following and preceding element because the following and preceding element are identical except that they are darker/brighter (in visual brightness), louder/quieter (in acoustical intensity), etc. Such as for music, in which the relationship (i.e., the distance) between one note and the other is important to create the melody, the present sequences, that were not musical, were created so the participant could perceive a relation among the various elements. It already known from literature that musicians can process the contour more automatically with respect to nonmusicians (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004b), but less is known concerning the ability of extracting a contour with nonmusical stimuli. A few studies observed that contour can be elicited also by nonmusical stimuli, such as loudness variations (McDermott, Lehr, & Oxenham, 2008; Prince, Schmuckler, & Thompson, 2009), but no previous study compared the perception of nonmusical contour of musicians and nonmusicians. The contour condition was therefore included for investigating whether there was a difference between people who underwent a music training and people who did not, in memorizing nonmusical elements with a contour structure. The nocontour stimuli, instead, had to be nonverbal (and not labelable) and non-musical (with no music characteristics). Details of how the contour and nocontour stimuli were created are described in the material section

Finally, we wanted to understand whether the short memory performance could be different in tasks in which the individual has to remember the serial order of the stimuli presented in a sequence, or in tasks in which s/he has to remember the single items presented in the sequence, regardless their order of presentation. In fact, some studies suggest that there are different processes for serial order memory and item memory (Gorin, Kowialiewski, & Majerus, 2016; Majerus, Poncelet, Greffe, & Van der Linden, 2006; Perez, Majerus, Mahot, & Poncelet, 2012). In sequence recall tasks (e.g., span forward), for instance, participants have always to reproduce the sequence in the correct order of presentation. Therefore, the order processing overlaps with the remembering of the items, and it is impossible to look at the two processes separately. Since serial order memory can be separated from item memory, it could be possible that the advantage of musicians previously observed in this kind of tasks is linked to only one (or mostly to one) of these two kinds of memorization. No previous study that compared the memory of musicians and nonmusicians controlled for these variables, therefore we believed it was important to assess whether the performance of musicians and nonmusicians could differ when remembering the serial order, or when remembering only the single items presented.

Note: in the present study we are still recruiting some subjects, to reach the planned numerosity. Here I will present the partial results obtained to this date.

Method

Participants

Thirty-six French native speakers participated to the study to this date. They were aged between 18 to 30 years old, had a normal to correct vision, had normal audition and did not have any kind of neurological or psychiatric disorder (assessed with self-report questionnaires). Demographic details of all participants are reported in Table 1. Participants were divided into two groups, 24 nonmusicians and 12 musicians (note that the study aims to collect a total of 24 musicians). Musicians were professional players or students at the music conservatory of Lyon and Villeurbanne, with a minimum training of 10 years. Nine musicians reported to have perfect pitch. Details of instrument played by musicians are represented in Table 2. All participants were payed for their participation, 10 or 12 euros per hour, depending on the protocol used.

Table 1. *Age, Education, Working Memory (WM) index and Speed of Processing index of the WAIS-IV, music education, and practice hours.*

	Musicians *N* =12 (8 females)	Nonmusicians *N* = 24 (19 females)		
Age (yrs)	26 (2.63)	23.7 (2.87)	***p* = .025**	*d* = .82
Education (yrs)	16.2 (2.34)	15.3 (1.02)	*p* = .645	*d* = .21
WM index	105.75 (9.69)	102.25 (12.97)	*p* = .371	*d* = .31
Speed of Processing index	111.83 (13.95)	108.75(14.18)	*p* = .540	*d* = .22
Music Training (yrs)	19.1 (4.29)	2.2 (0.96)		
Weekly Practice (hrs)	24.8 (14.21)	/		

Note. Mean (SD). The last two columns to the left represent the p-value of the Welch t-test used to compare the two groups and the Cohen's *d*. WM index is the standardized score of the two Working Memory measures subtests used, the Digit Span and the Arithmetic. Speed of Processing index is the standardized score of the two WAIS-IV subtests, the Code and the Symbol.

Table 2. *Instrument played by each musician.*

Subject	Instrument
1	Viol (Viola da Gamba)
2	Cello
3	Cello
4	Bass Trombone
5	Harp
6	Cello
7	Violin
8	Violin
9	Violin
10	Trumpet
11	Viola
12	Clarinet

Note. Several musicians were playing more than one instrument, here only the main one was reported

General paradigm

A recognition paradigm was chosen for the short-term memory task, because it could be easily adapted with the different type of stimuli used described under. The task required to see/hear a first sequence of elements (from now on, S1), and to compare it with a successive sequence of elements (from now on, S2). After the presentation of S2, the subject had to judge by clicking on the mouse if S2 was the same of S1 or if it was different. When S2 was different from S1, S2 could be different in two ways: the order of the elements of S1 could be changed or there could be a new element that replaced one of the elements of S1. From now on, I will refer to the different S2 sequences with the order change as “order condition” and with the new element as “item condition”. In the order condition, to create S2 only adjacent items of S1 were inverted, and equally in every position of the sequence (e.g., ¼ of the changes were made by switching position 1 and 2, ¼ by switching position 2 and 3, and so on with 3 and 4, and 4 and 5). In the item condition, the position in which the new item replaced one of the items of S1 was equally distributed (e.g., 1/5 of the changes were made in position 1, 1/5 in position 2, 1/5 in position 3 and so on). When creating the S2 different sequences, we were making sure that the contour (e.g., relationship between items, see “material” section for details) was always changing. For example, a luminance S1 and S2 could be the following: “first stimulus; “second stimulus darker” “third stimulus brighter” “fourth stimulus darker” “fifth stimulus darker”, the new sequence should be “first stimulus” “second stimulus brighter” “third stimulus brighter” “fourth stimulus darker” fifth stimulus darker”. We opted for changing always the contour because the stimuli were already abstract, and for this reason not easy to remember. The sequences were presented via the software Presentation. An example of the paradigm is represented in Figure 1.

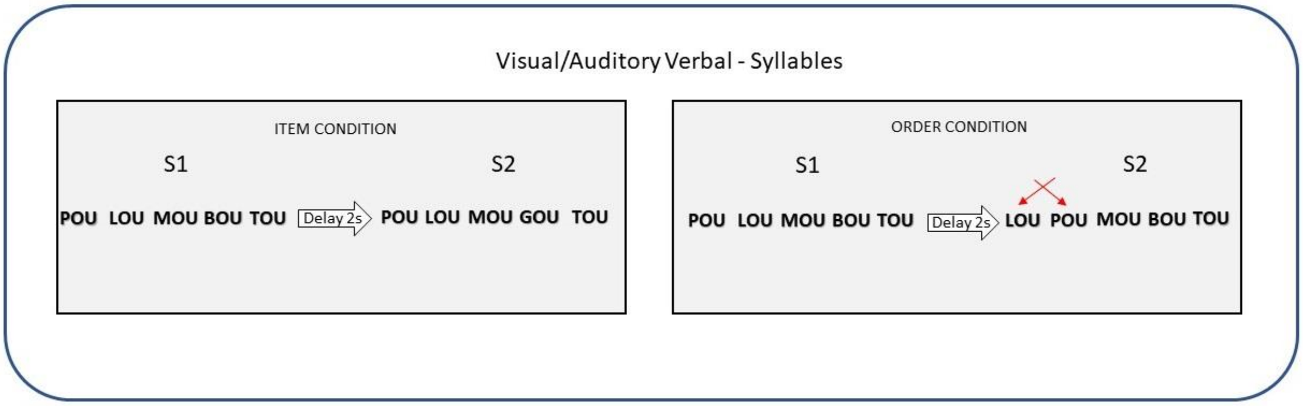

Figure 1. Examples of verbal trials with different S2: on the left, an item condition trial, in which one item was different of S2 was new with respect to S1; on the right, an order condition trial, in which two elements of S1 were inverted in position in S2. In the visual presentation the syllables appeared one at a time from left to right, in the auditory presentation the same syllables were played via headphones with the exact same duration as for the visual presentation.

Material

All the stimuli here described were chosen, created, and modified after being tested with a pilot study. They will be classified in two ways: firstly, depending on the modality of presentation, that is, stimuli presented auditorily, and stimuli presented visually, and secondly, depending on the type of stimuli presented: verbal stimuli, stimuli with contour, and stimuli without contour. Contour stimuli included a pattern of elements that could be perceived and remembered by the relation across elements (e.g., “up-down-up-up-up”). Within each category of contour stimuli, the elements presented belonged all to the same category, but they changed with respect of one characteristic. For example, contour is usually experienced in music. When hearing a melody, what makes a melody is the relationship across different notes (e.g., the second note is higher than the first, and it creates a feeling of "ascension"). The various relationships between elements define the "contour" of the melody. We opted to recreate the same feeling with nonmusical stimuli, by changing one characteristic of our stimuli: In the case of the auditory stimuli, we manipulated the loudness (i.e., louder vs quieter), since previous studies reported that contour could be experienced also with changes in loudness (McDermott et al., 2008; Prince et al., 2009); in the case of visual stimuli, we opted for changing the luminance (i.e., darker vs brighter) basing on the crossmodal correspondence with the loudness stimuli (Marks, 1987).

Stimuli presented auditorily

AUDITORY VERBAL: in this condition, we chose to have a pool of six different syllables from which to create five-item sequences. The syllables were: "pou", "mou", "bou", "lou", "gou", "tou". These syllables were previously created and used by Tillmann and colleagues, (2009), and then adapted for the present study. Each syllable had a fundamental frequency of 232 Hz, and they had been equalized in terms of intensity. Intensity was also checked with a phonometer, in order to be sure that they were presented at ~67 dB SPL. Each syllable was 500 ms long. The inter-stimulus interval (i.e., ISI) was 100 ms long and the pause between S1 and S2 was 2000 ms long. See Figure 1.

AUDITORY CONTOUR: a pool of six different tones was used to create a contour with nonmusical stimuli. Sequences of five-items were created. The tone used was a piano note, the A#2, therefore its frequency was the same as for the syllables (i.e., 232 Hz). The software Presentation was calibrated in order to present the tones at the following dB SPL: ~55 dB for the softer tone and ~80 dB for the loudest tone, and each tone between these two levels increases (or decreases) in steps of 5 dB. dB levels were also checked with a phonometer. Each tone lasted 500ms, the ISI was 100ms and the pause between S1 and S2 was 2000ms long. See Figure 2 for an example of the trials.

AUDITORY NOCONTOUR: In this category, a pool of five pink noises was used to create a set of four-items sequences. The length of the sequence was shorter than that of other materials because of the difficulty emerged in the pilot study and because of the technical limits of creating different sounds without changing the pitch. The stimuli were created in a collaboration with the IRCAM institute of Paris. Five pink noises were created with the software Cool Edit Pro, and then modified with the software Cecilia and its "particle" module. This manipulation allowed to change the texture of the noise with a granulator synthesis method (Roads, 1988), and to create noises that can be discriminable without contour variations (note that an impression of contour with these noises was still reported by some subjects). The loudness of the noises was checked with a phonometer in order to be presented at

~67 dB SPL. Each noise was 500 ms long, the ISI was 100 ms and the pause between S1 and S2 was 2000 ms long. See Figure 2 for an example of the trials.

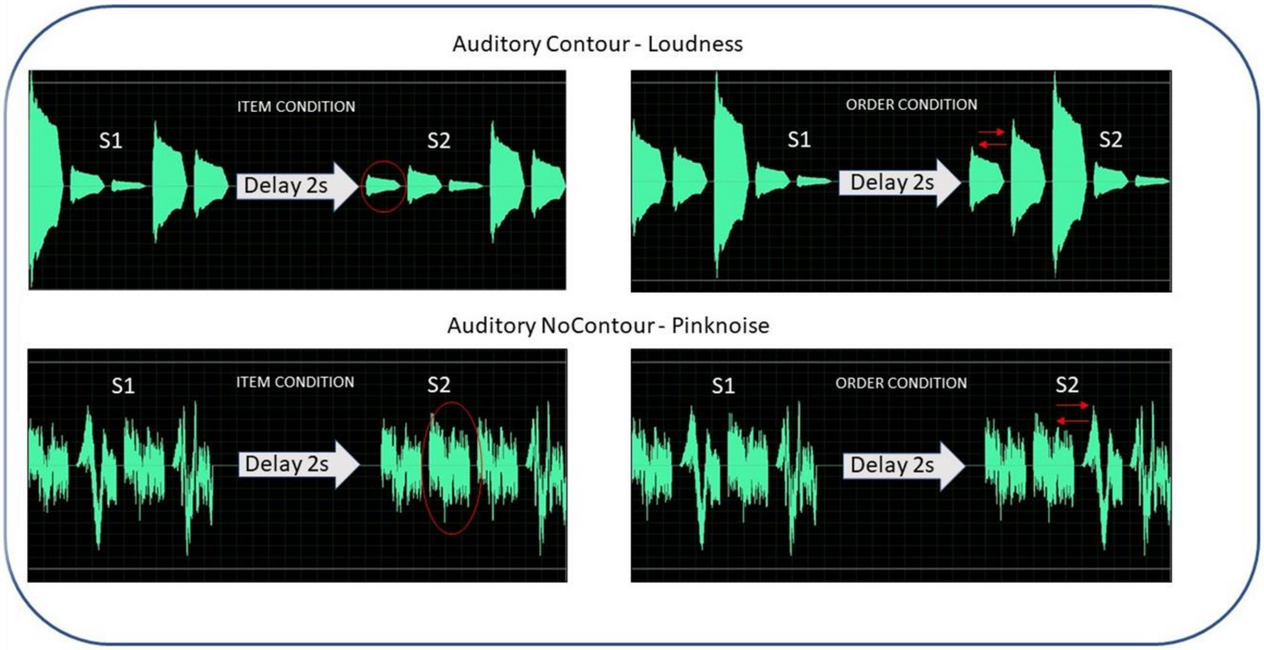

Figure 2. Examples of auditory trials: on the top, the waveform of the contour items, which were changing only in intensity levels (i.e., dB). On the bottom, nocontour items, made of pink noises that were changing in texture. On the left, example of item condition trials, on the right, example of order condition trials.

Stimuli presented visually

VISUAL VERBAL: in this condition, the exact same six syllables presented in the auditory modality were used. The syllables were presented in black color on a white background, and they were presented in five different positions on the horizontal axis of the computer screen, from left to right, and always centrally in the vertical axis. The timing was the same as for the auditory condition. See Figure 1.

VISUAL CONTOUR: a pool of six squares with different grades of luminance was selected. Sequences of five-items were then created. The square was presented at the centre of the screen, on a black background, and each level of luminance increased/decreased in steps of 5 cd/m2: The dimmest level had a luminance of 2.26 cd/m2. Each item lasted on the screen for 600 ms and there was no ISI between items. This choice was made because a pause of 100 ms between the items, produced a flickering effect that disturbed the perception of the various items. The pause between S1 and S2 was

2000 ms. Example of sequences are represented in Appendix 2. Note that a similar task was already presented in a couple of studies, (Aizenman, Gold, & Sekuler, 2017; Gold, Aizenman, Bond, & Sekuler, 2014), but with different purposes. See Figure 3 for an example of the trials.

VISUAL NOCONTOUR: for this category, we decided to have a pool of five different Japanese ideograms (i.e., kanji) from which to create four-item sequences. Also in this condition, the sequence was shortened because having items that do not have an associated semantic meaning, at the same pace as for the other stimuli, made the task too difficult. The number of elements (i.e., horizontal, vertical, and oblique lines) in each item was equalized, in order to avoid any coding related to this variable. In this way, some elements of the ideograms were cut, and some other were added with the software Inkscape. Each kanji ideogram was presented for 500 ms, followed by a 100 ms of ISI. The pause between S1 and S2 was 2000 ms. See Figure 3 for an example of the trials.

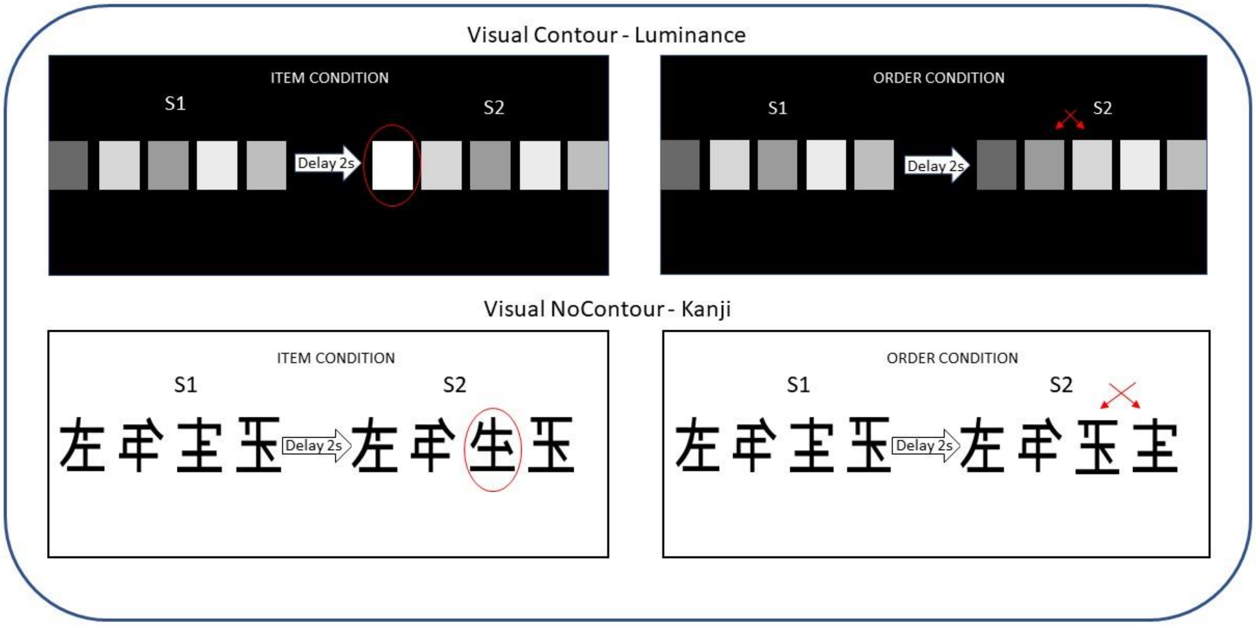

Figure 3. Examples of visual trials: on the top, the contour items, which were changing only in luminance levels. On the bottom, nocontour items, composed by Japanese Kanji ideograms. On the left, example of item condition trials, on the right, example of order condition trials.

WAIS-IV (Wechsler Adult Intelligent Scale, Wechsler, 2008) – Working Memory subtests.

Digit Span: this subtest is composed by three tasks. In the first one, the digit span forward, the participant is asked to repeat a list of digits that is read by the experimenter, in the same order of presentation. There are two trials with the same length, and then the list increases of one item. When the subject makes two mistakes in two trials of the same length, the task is interrupted. The second task is the digit span backward. It follows the same principles of the span forward, but the difference is that the subject is required to repeat the digits in the reverse order with respect to their presentation (i.e., from the last digit to the first one presented). The final task is the digit ordering span, in which the subject has to repeat the digits presented starting from the smallest number and proceeding in ascending order. Each correct response scores one point, and the total scores is the sum of correct responses.

Arithmetic: in this subtest, the subject hears a series of arithmetic problems that are read by the experimenter, and s/he has to solve mental operations starting from easy ones and proceeding with more difficult ones. When the subject does three consecutive mistakes, the task is interrupted. Each correct response scores one point, and the total scores is the sum of correct responses.

WAIS-IV Speed of processing subtests

Symbol: in this task the participant has to compare whether two symbols depicted on the left side of a paper are present in a list of symbols on the right side of the paper. If one of the two symbols is present, the subjects have to mark it, otherwise, s/he has to mark “no”. There are 10 trials per page, and six pages in total, and the subject has to perform the task as fast as possible, within two minutes. The score is calculated by adding up all correct responses and by subtracting all incorrect responses.

Code: in this timed task, the participant sees a box with a series of numbers, each one associated to a nonsense symbol. There are ten numbers and ten corresponding symbols. The participant has then to fill up the empty boxes below each number of a series, with the respective symbol as indicated in the reference box. This has to be done the fastest as possible, and after two minutes, s/he has to stop. The score is calculated by counting all the symbols correctly reported in the boxes.

Questionnaires: some questionnaires were given to the participant to collect demographic details such as, age, sex, education, profession, handedness, presence of hearing or vision problems. Then specific questions about music training were administered to participants, such as if they played, singed or danced, for how long, starting age, practice hours, if they underwent specific trainings, if they had perfect pitch and some feelings about their ability and about music in general. Finally, a questionnaire about the experience was administered, to collect opinions about the difficulty of the tasks and the strategies used.

Procedure

The participant started by filling up the questionnaires about the demographic details and the music training. They then moved to a silent small room to perform the computer tasks. As first thing an audiometry was run, to check if they had a normal audition. In order to be included in the study, participants had to have at least a threshold of 20 dB HL for 250, 500, 1000, 2000, 4000, and 6000 Hz. Secondly, all the instructions for the short-term memory task were given, and the stimuli were showed (and played, in the case of auditory stimuli) to the participant on a power point presentation. The participant then completed a training session before each of the six blocks corresponding to the six different types of stimuli. In the training session there were 50% of same trials (i.e., S1 = S2) and 50% of different trials (i.e., S1 ≠ S2). Successively, the experimental block was composed by 40 trials (for the verbal and contour stimuli) or by 48 trials (for the nocontour stimuli). In each experimental block, there were 50% of same trials and 50% of different trials. Participants were doing six blocks with always the same kind of difference in S2 (i.e., item condition, order condition). The first part lasted around 45 minutes. After all the six blocks, the WAIS-IV subtests were administered. The duration of this second part was about 30 minutes. Finally, subjects were doing the third and last part that consisted in doing the same six blocks (in the same order as for the first part), but with the other type of S2 difference (e.g., item condition if they started with the order condition, and vice versa). Half of subjects were starting with the order condition, and the other half with the item condition. The order of the blocks was also counterbalanced across subjects, and also the structure of the trials within each block (i.e., the items that were composing the sequences) was changing across subjects. The counterbalancing avoids having the same pattern of trials (e.g., position of change) across materials, and plus, moreover, there were not more than three trials in a row with the same answer (i.e., three consecutive trials with only "same" answers or with only "different" answers). At the end of the third part, the questionnaire about the experience was administered. The duration of the whole experimental session was about three hours.

Analysis

First of all, groups were compared in terms of age, education, and WAIS-IV subtests (see Table 1), with Welch's t-test, which is preferable when the numerosity and the variance (as indicated by Levene's test) of the groups are unequal. To analyze the short-term memory task results, firstly the percentage of correct responses was calculated separately for same trials (i.e., S1 = S2) and for different trials (i.e., S1 $\neq$ S2), and hits (number of correct answers for different trials) and false alarms (number of incorrect answers for same trials) were calculated. After having these data, *d'* was calculated, to have a measure of the ability of the subject to detect a difference between S1 and S2. When *d'* is statistically different from zero, it suggests that the performance was not at a chance level. Moreover, the higher the value, the better the performance. All the statistical analysis on the subjects' performance in the memory task were based on *d'*. The criterion *c* was also calculated, because it allows to have an estimation of the response bias. Positive *c* values (and statistically different from zero), reflect the tendency of the subject to respond "same" to the trials. This could suggest that the subject had difficulties in detecting the different trials. On the other side, negative values of *c* (statistically different from zero) reflect the tendency of answering always "different". In this case the subject would "see" or "hear" differences S2 when in reality they are the same as S1. *c* and *d'* were used as dependent variable in the following repeated measures ANOVAs. As within variables, the following factors were included: modality (i.e., visual and auditory presentation), condition (i.e., item and order), type, which refers to the type of elements of each block (i.e., verbal, contour, nocontour); the group (i.e., musicians and nonmusicians) was included as between factor. Welch's t-test was then used in post-hoc analyses to investigate the significant effects of the ANOVAs. The statistics of significant results of ANOVAs will be always reported, whereas in the case of nonsignificant results, only the main nonsignificant effects will be reported. In post-hoc tests, the statistics of nonsignificant results will be not reported, to ease the reading. The p-values for multiple comparisons here reported are all FDR adjusted. The analyses were performed with the software R and JASP.

Results

Participants were comparable in terms of education and of WAIS-IV subtests, but they were significantly different in age (see Table 1). Nevertheless, age was not included as a covariate because the mean difference is anyway small (i.e., 22 vs 24), and memory performance is not considered to vary in this age range. The first analysis was run to check whether *d'* values were significantly different from zero, separately for each one of the six types of stimuli, and separately for item and order condition (i.e., the 12 blocks presented to participants). All our *d'* were statistically significant from zero, for all the 12 blocks (always $p < .001$, FDR corrected). The same analysis was also run separately for groups (because the performance of the two groups was different), and again all the *d'* values were significantly different from zero (always $p < .001$, FDR corrected). *d'* was then included in a first repeated measure ANOVA, with group as between factor, and modality, condition and type as within factors. Given the large number of variables included, each effect will be described separately for each variable.

Condition. The condition (i.e., order, item) was not significant, F (1,34) = .14, $p = .709$, $\eta^2 = .004$, meaning that participants performed equally when there were item changes in S2 ($M = 1.82$, $SD = 1.01$) and when there were order changes in S2 ($M = 1.83$, $SD = .99$). Qualitatively, most of participants reported to be not aware of the difference between order and item condition, even if they were told before the task which condition they would have performed. There was a significant interaction between condition and type of stimuli, F (2,68) = 6.26, $p = .003$, $\eta^2 = .15$, but post-hoc tests did not evidence any significant difference between item and order condition separately for each type of stimuli, and the largest difference between conditions was found with nocontour stimuli (mean difference in d': 0.20, $p = .071$). This could reflect a low statistical power. However, there was a significant interaction between condition, modality, and type, F (2,68) = 11.99, $p <.001$, $\eta^2 = .25$. Post-hoc tests revealed that there was a significant difference between item and order condition only with auditory nocontour stimuli, t (69) = -3.29, $p = .009$, where the performance was higher in the order condition ($M = 2.06$, $SD = .93$) with respect to the item condition ($M = 1.37$, $SD = .82$).

Modality. The modality (i.e., visual, auditory) was significant, F (1,34) = 12.01, $p =.001$, $\eta^2 = .25$, in fact, the performance of all participants was slightly better with the stimuli presented visually ($M = 1.99$, $SD = 1.05$) than with the stimuli presented auditorily ($M = 1.66$, $SD = .91$). A significant interaction between modality and type was found, F (2,68) = 6.98, $p = .002$, $\eta^2 = 15$. Post-hoc tests revealed that there was a difference between modalities only with the verbal stimuli, t (140) = 4.37, $p <.001$, where the performance was higher with syllables presented visually ($M = 2.57$, $SD = .90$) with respect to syllables presented auditorily ($M =1.94$, $SD = .81$) There was no interaction between modality and group, F (1,34) = 1.46, $p = .235$, $\eta^2 = .03$: qualitatively, musicians performed slightly better than nonmusicians with auditory stimuli (musicians: $M = 2.11$, $SD =.86$; nonmusicians: $M = 1.44$, $SD = .86$), and with visual stimuli (musicians: $M =2.30$, $SD = .99$; nonmusicians: $M = 1.83$, $SD = 1.05$), but not particularly in only one modality.

Type: the type of stimuli (i.e., verbal, contour, nocontour) was significant, F (2,68) = 16.43, $p < .001$, $\eta^2 = .25$, suggesting that performance of all subjects vary depending on the type of stimuli presented. Post-hoc analysis, revealed that performance of all subjects was higher with verbal stimuli ($M = 2.26$, $SD = .91$) than both contour stimuli ($M = 1.48$, $SD = 1.06$), t (279) = 6.64, $p < .001$ and nocontour stimuli ($M = 1.74$, $SD = .85$), t (285) = 4.96, $p < .001$. Performance with contour stimuli was also significantly different from the one with nocontour stimuli, but here the difference was much smaller than the two previous ones, t (273) = -2.3, p = .022. There was an interaction between group and type (see fig 2 and fig 3), F (2,68) = 14.24, p < .001, η^2 = .22, and post hoc tests suggested that musicians performed better than nonmusicians both with contour stimuli (where the difference between groups is large), t(79) = 6.88, p < .001, and with nocontour stimuli (where the difference between groups is still large, but smaller than for the contour stimuli), t (88)= 3.17, p = .003. Interestingly, with verbal material, there was no statistically significant difference between groups, t (96) = .34, p = .736. There was also an interaction between group, type, and modality (see Figure 2 and Figure 3), F (2,68) = 5.79, p = .005, η^2 = .12. Specifically, post-hoc tests revealed that with the visual modality (see fig 2), musicians outperformed nonmusicians with contour stimuli, t (38) = -5.04, p < .001, where the d' of musicians was more than the double of the one of nonmusicians, but no difference between groups was found with nocontour stimuli, t (52) = -.09, p = .998, and with verbal stimuli, t(48) = .02, p = .998. With the auditory modality (see fig 3), musicians outperformed nonmusicians with contour stimuli, t(44) = -5.07, p < .001, and on average, musicians had one point more in their d' than nonmusicians; musicians still outperformed nonmusicians with nocontour stimuli, t(39) = -4.16, p < .001, and again, the d' of musicians was almost one point more than nonmusicians; finally, with verbal stimuli, no significant difference between musicians and nonmusicians was found, t (45) = -0.53, p = .901, and there was only a small difference in d' between groups, with musicians performing slightly better.

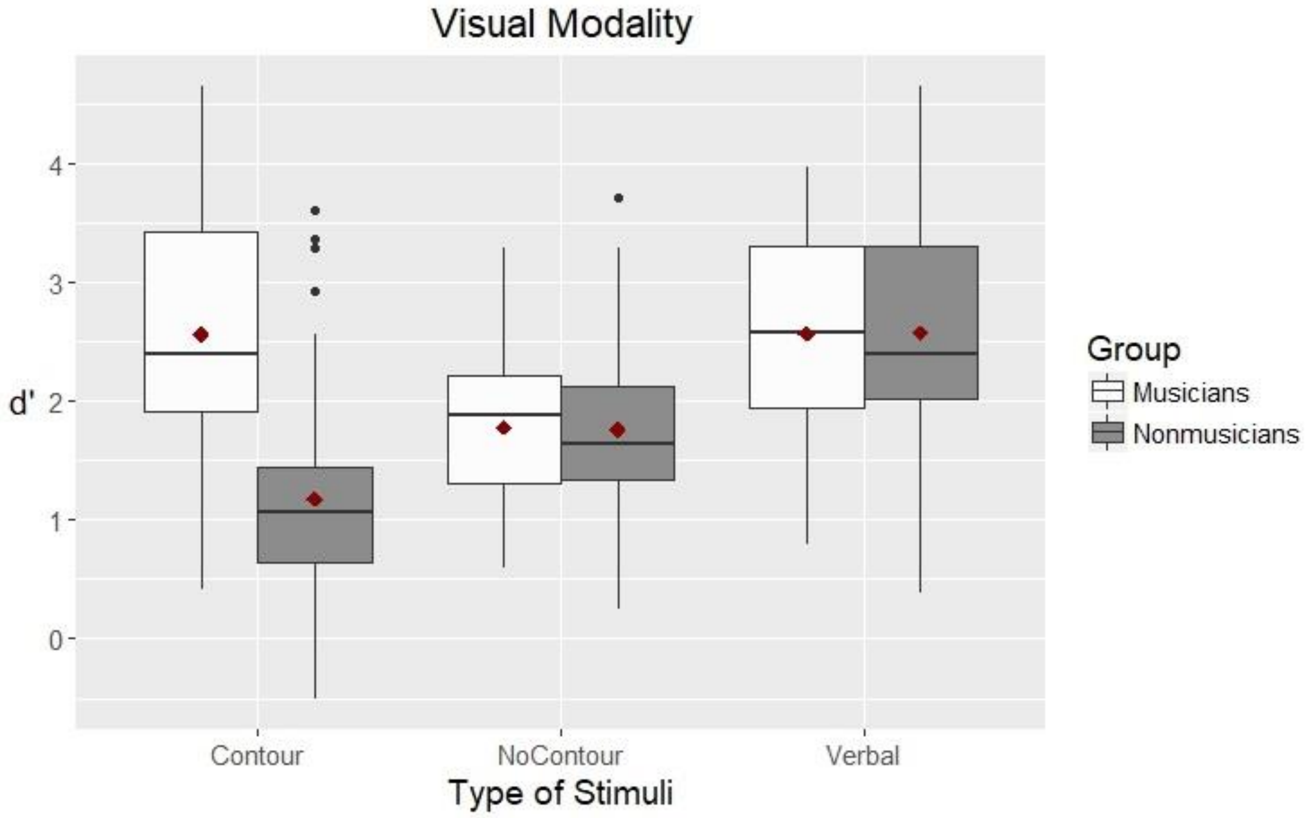

Figure 2. Accuracy (*d'*) for the memory task, visual modality only, separately for each type of stimuli and group. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The black dots are outliers.

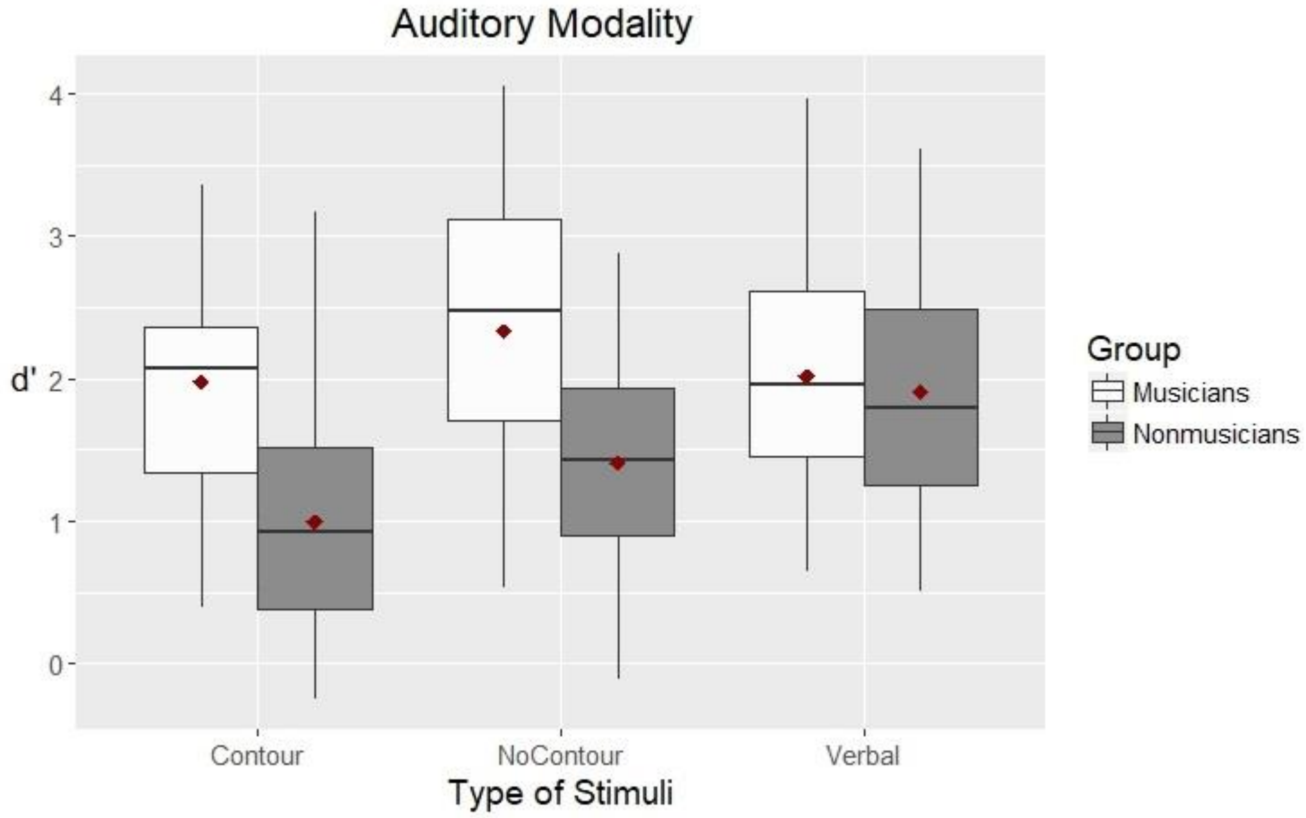

Figure 3. Accuracy (*d'*) for the memory task, auditory modality only, separately for each type of stimuli and group. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%.

Criterion *c*. To assess the presence of a response bias, c was calculated. Firstly, t-tests were run for each one of the 12 blocks, to check whether there were values significantly different from zero, which indicate a bias towards "same" answers (i.e., positive values) or towards "different" answers (i.e., negative values). T-tests Bonferroni corrected revealed that with contour and nocontour stimuli, in each condition and modality, the c value was statistically different from zero, (p <.001), revealing a response bias towards "same" answers. A repeated measures ANOVA was then run with condition, modality, and type as within factors and group as between factor. The condition was not significant, F (1,34) = 2.88, p = .099, η^2 = .07, meaning that there was no difference in the response bias between item and order condition. There was not an effect of modality, F (1,34) = 3.39, p = .074, η^2 = .09. A significant effect of type of stimuli emerged, F (2,68) = 42.52, p < .001, η^2 = .55., see Figure 4. Post-hoc comparisons showed that there was a significant difference in *c* value between verbal and contour stimuli, *t* (277) = 7.80, *p* <.001, with only contour stimuli associated to a response bias towards "same" answers. Another difference emerged between verbal and nocontour stimuli, *t* (270) = 9.79, *p* < .001, where nocontour stimuli were also associated with a response bias towards "same" answers. A small difference was found also between contour and nocontour stimuli, *t* (285) = 1.99, *p* = .047, and the response bias was slightly higher with the nocontour stimuli. There was not an effect of the group, *F* (1,34) = .22, *p* = .644, η^2 = .01.

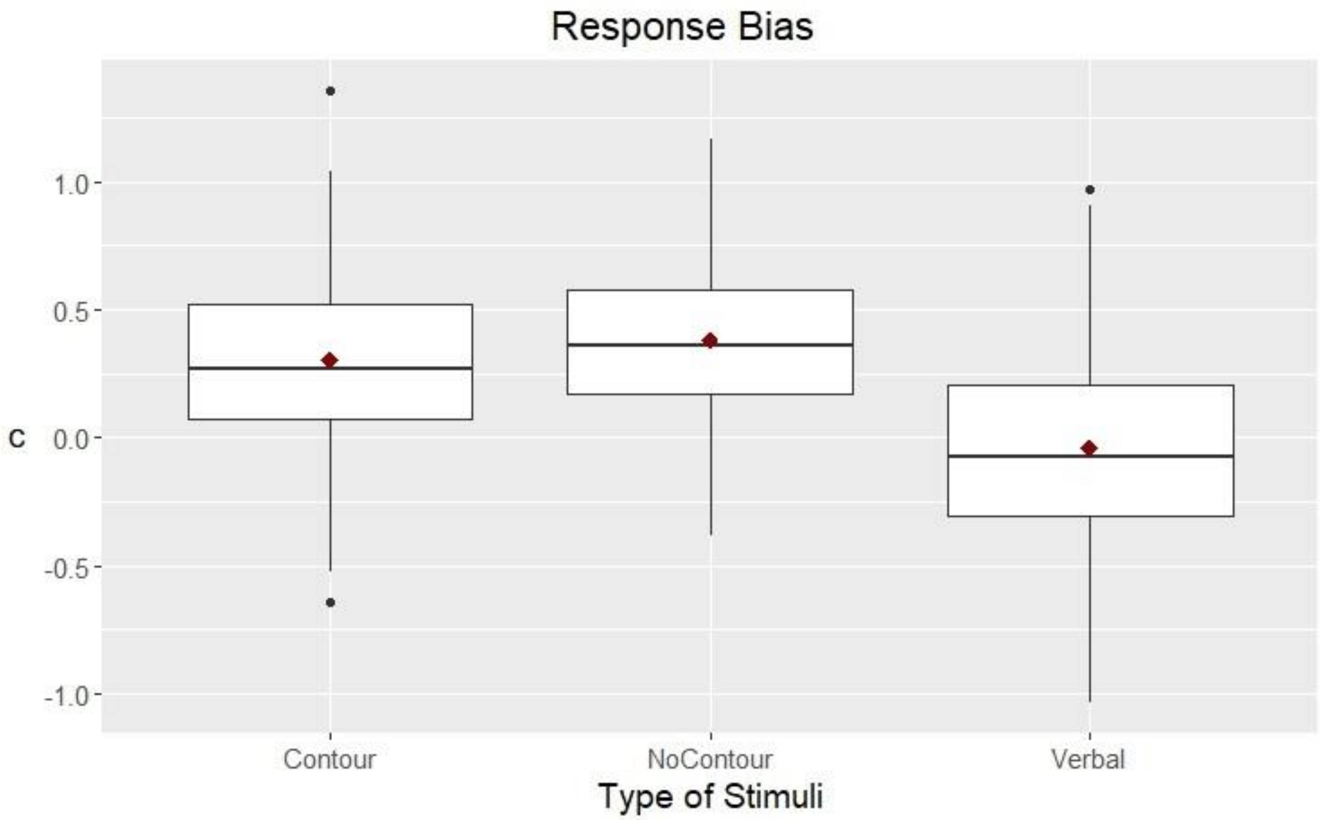

Fig 4. Response bias (*c*)for the memory task, separately for each type of stimuli. Inside each box, the horizontal line indicates the median. The red diamond represents the mean. The edges of the box represent the 25th and the 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50%. The black dots are outliers.

Discussion

The present study investigates the performance of musicians and nonmusicians on a short-term memory task, for different types of materials: verbal stimuli, and nonmusical - nonverbal stimuli. In fact, no previous study that compared the memory of musicians and nonmusicians had ever presented nonverbal and nonmusical stimuli, both in the visual and the auditory modality. In the present study, the aim was to see whether musicians could perform better not only with auditory verbal stimuli, as suggested by literature (see STUDY 1), but also with auditory nonverbal and nonmusical stimuli. For these stimuli, a contour condition and a nocontour condition were included, to see whether the advantage of musicians extended to stimuli in which there is no music but there is a characteristic that belong to music (i.e., contour), and with stimuli that are not verbal neither musical. Moreover, the modality of presentation was also considered, and the stimuli could be either visually presented, either auditorily presented. Finally, since literature suggested the importance of distinguishing between the memory for serial order and the memory for the items (e.g., Majerus & Cowan, 2016), two different conditions were included, one with serial order changes, and one with item changes. We expected a superior performance of musicians in auditory memory, regardless the type of stimuli presented.

The partial results here described showed that there was not a main effect of group, therefore musicians were not overall better than nonmusicians in this task. There was not any modality effect (i.e., different performance for visual stimuli and auditory stimuli) and no interaction between modality and group, contrarily to what expected. However, a significant interaction between the group and the type of stimuli emerged, and the dimension of the effect was moderate. Interestingly, musicians performed better than nonmusicians with the contour and nocontour stimuli, but no difference in the verbal condition was found. Moreover, another interaction between group, type of stimuli, and modality emerged, showing that the advantage of musicians was more pronounced with auditory contour and nocontour stimuli, and with visual contour stimuli. This suggests that musicians benefit from the contour condition, independently from the modality of presentation, and the difference between groups in these conditions was quite large. Possibly, their expertise allows them to easily extract the contour of a sequence of stimuli, even if this sequence is not musical, and to remember it as a unique structure instead that of single separated elements. After the experiment, many musicians reported they could perceive a “melody” while performing the visual contour task, and they were associating, more or less automatically, an imagined sound to a specific level of luminance. Note that this association between the luminance sequences and the imagined melody was already reported by some subjects in the study of Gold and colleagues (2014), but they did not include a group of musicians, nor controlled for music training (Gold et al., 2014). With the auditory contour stimuli, participants reported to perceive the sequences of loudness also as a melody, or as having a rhythm. It is already known from the literature that loudness changes can induce a feeling of a changing in pitch, even the frequency if tones is fixed (McDermott et al., 2008). However, in the present study only musicians seemed to perceive musical features with these intensity variations. It is interesting to note, that this feeling of pitch changes can extend to visual stimuli too (i.e., the luminance changes), especially if the person who carries on the task has a music expertise. The way musicians reported having this feeling reminds of synaesthesia phenomena (for a review, see Martino & Marks, 2001), and it appears to be mainly automatic (i.e., the musician did not decide to use that particular strategy, but reported to have the automatic association with music). Nonmusicians, on the contrary, performed quite poorly in the contour conditions, both in the auditory and the visual modalities, and only few of them were able to use some strategies (such as imagining/hearing a melody, as done by most of the musicians), therefore they had troubles in remembering the sequence because, for them, the elements were too similar to one another.

Concerning the nocontour stimuli, here musicians performed better than nonmusicians in the auditory modality in particular (i.e., the pink noise sequences). This result could have two explanations: one is that musicians exploit their fine auditory perception skills, and therefore can better distinguish the various items with respect to nonmusicians, who could presumably find it more difficult to distinguish the various noises. Another possibility is that, instead (or also), musicians could perceive the different sounds as having different pitches. Even though the pink noises selected for the present study did not have a specific frequency as pure tones have, it is possible that the different texture of the noises could elicit different pitch perception. Some musicians reported, in fact, that they could hear a contour, even with the pink noise stimuli. This could have helped them in remembering the sequences as having, again, a contour. This last case was not the one that was hoped for, because the aim was to have an auditory condition that could not elicit any verbal labels and any musical feeling. However, it is possible that this is a limit of the world of sounds, and that people with a music background would easily hear music features even in nonmusical stimuli.

The fact that here there was no difference between groups with the verbal sequences, especially with the auditory presentation, was not expected, since literature suggests that usually musicians performed better than nonmusicians with verbal memory tasks, and mostly when the task is presented auditorily (e.g., Franklin et al., 2008; Tierney et al., 2009, STUDY 1, STUDY 2). This can have different explanations. One is that the task could be too simple for all participants, especially the visual one, in which most of participants reported to read and memorize only the first letter of the syllables. The performance was much higher for both groups with respect to the other type of stimuli, as also reflected from the significant effect of type in the ANOVA, suggesting that the level of difficulty might have hidden some eventual differences between groups. Another possibility is that this particular task was not often used in investigating verbal memory of musicians and nonmusicians. If this was the case, a problem of defining which task investigates what in memory would emerge: also STUDY 2 showed that there is a moderate variability among studies, and also among the tests that are chosen to tap a specific memory aspect.

Concerning the condition, here no difference between the item and the order condition was found, even though previous studies found often that the performance was easier in the item condition than in the order one (Majerus et al., 2006; Martinez Perez, Majerus, & Poncelet, 2012). It could be possible that the recognition paradigm is not the best option to investigate these two different types of memorization. To test memory for serial order, often recall paradigm was used, and the order condition was assessed by giving, for example, some cards with the items presented before pictured on them, and the participant had to place each card (i.e., item) in the right order of presentation (Martinez Perez, Majerus, & Poncelet, 2013). Here, however, it was impossible to use a recall paradigm with the auditory stimuli.

Finally, regarding the response bias, the analysis of the criterion *c* showed a bias towards the answer "same" for contour and nocontour stimuli, in both modalities. This suggests that in those blocks, different sequences were particularly difficult to detect. In fact, this bias is more likely associated to the difficulty of those specific tasks than to a general response tendency. In fact, if the bias was general, participants would have shown it also when performing the verbal blocks. Subjective reports from the participants confirmed that the blocks with contour and nocontour stimuli were perceived as being particularly difficult. Nevertheless, there was a large variability in judging which one among all the type of stimuli was the hardest, and this suggests that short-term memory performance largely depends also to subjective characteristics, such as strategies used, sensitivity to sounds, sensitivity to image details, etc.

Some limitations of the present study have to be acknowledged. The first is the number of items per sequence: the initial idea was to have the same number of elements per sequence for each type of stimuli, to make the comparison among different stimuli more reliable. For nocontour stimuli, the sequence had to be reduced from five to four stimuli because of the difficulty of (1) remembering five items sequence where no verbal labels are possible, and no musical features are present, and (2) because the creation of the items, especially for the noises, was technically difficult, because any further modification of the noise produced a feeling of different pitch, or it was too similar to the others; in other words, creating even one more noise would have led to a sequence in which either there was a clear impression of pitch changing, or either some elements were too similar to each other to be distinguished. Another limitation of the nocontour stimuli is that the aim of having auditory and visual stimuli that could not elicit any verbal label (otherwise the person would have remembered a sequence of labels, for example), and any musical feeling (such as contour), was difficult to accomplish. In fact, when the task is particularly difficult, subjects try to use different strategies and tend to give a meaning to what they see/hear. This meaning could be accompanied by a verbal label, as reported by some participants. Few of them, in fact, reported, as strategy used, to attach a label to the different pink noises, even if they were really similar to each other. The strategy was not associated to a particularly high performance anyway, but this suggests that it is rather difficult to create stimuli in which the person does not find any label. Moreover, with the pink noises some musicians reported to feel a contour, and this because they probably tend to hear more subtle differences among sounds. Nevertheless, this limitation is difficult to avoid, because if the sounds are too similar to each other, the sequence become almost undistinguishable in the memory task.

Regardless these limitations, the present study (to the date), reveals an interesting fact: musicians can extract a meaningful structure (i.e., the contour), also in nonmusical sequences of stimuli. This suggests that their training might help them developing strategies for learning and memorizing items. Concerning auditory stimuli, the advantage might be driven by a general superior discrimination of sounds (Strait & Kraus, 2014), but interestingly, musicians can also use their expertise to perform better in visual tasks where the stimuli are in a relationship with each other. Of course, the present study is still ongoing, therefore any effect here described is not definitive. However, a clear tendency emerged already.

Acknowledgements

I would like to thank Patrick Bouchet of the DYCOG team (CRNL, Lyon) for programming the memory task, and Olivier Houix of the IRCAM institute of Paris for helping me in designing the nocontour auditory stimuli.

CHAPTER 5. GENERAL DISCUSSION

5.1 Synthesis of results

The present dissertation focused on the relationship between music training and memory skills. In particular, I investigated the short-term memory and working memory of musicians in comparison to nonmusicians. Literature on experts shows that high skilled people differ from less-skilled people not only in their area of expertise, but also in terms of brain activity, brain anatomy, and cognitive functions (Jäncke, 2009; Maguire et al., 2000; Reingold, Charness, Pomplun, & Stam, 2016). Expertise is known to affect the brain and the behavior, and musicians are a good model to study the long term effect of expertise (Münte, Altenmüller, & Jäncke, 2002). In fact, musicians, to become so, have to undergo a long and intense training, which requires commitment, motivation, but, mostly, a lot of practice. Several studies suggested that the music training benefits the development of cognitive functions (for a review, Schellenberg & Weiss, 2013). However, there is not yet clear evidence of a causal relationship between, for instance, music training and enhanced cognitive skills, because of a lack of true experiments. As reviewed in the first chapters, most of studies that focused on musicians are quasi experiments. Therefore, it is difficult to tell whether the differences observed between musicians and nonmusicians are caused by the music training or if they were pre-existing individual differences.

Here, I reviewed how musicians possess enhanced cognitive skills with respect to nonmusicians, in particular memory skills. Starting from the fact that musicians seem to have better memory skills than nonmusicians (e.g., Franklin et al., 2008; Hansen et al., 2013; Roden et al., 2012), the aim was first to clarify this advantage. STUDY 1 investigated the literature on memory skills of musicians and nonmusicians, through a meta-analysis. The aim was to understand whether there is true difference between groups in memory tasks, and how large this difference is. Literature was analyzed separately for long-term, short-term, and working memory. A moderate effect size, which suggests that there is a difference between groups, of medium dimension, was found in both working memory and short-term memory tasks. In short-term memory a small evidence of a publication bias emerged, but even when controlling for it, the effect size remained moderate. In long-term memory, a small effect size emerged, and plus, when controlling for a possible publication bias, the effect size was reduced and became no longer significant. The meta-analysis took into account also the different type of stimuli that were used in the various tasks (i.e., verbal, tonal, and visuospatial), because often they are thought to be differently processed and stored (e.g., Baddeley & Hitch, 1974). The type of stimuli was included as a moderator, which allows to see whether the effect size observed varies depending on the stimuli presented in the task. The effect-size was indeed different among the three categories: it was always large for tonal stimuli (i.e., the difference between musicians and nonmusicians in memory tasks with tonal stimuli is large), moderate for verbal stimuli (i.e., there is moderate difference between groups in verbal tasks), and from moderate to small for visuospatial stimuli (i.e., the difference between groups is moderate in visuospatial short-term memory tasks and small in visuospatial working memory tasks). This difference suggested that there is a variability across studies' results, but that that it depends partly on what has to be remembered (i.e., the type of stimuli). Interestingly, the meta-analysis showed that musicians performed better than nonmusicians with verbal stimuli too, and this advantage is quite consistent across different studies and tasks. It could be a case of generalization of the advantage (transfer from music skills to verbal skills), but it could be also that music and language share common processes. The debate is still open (Caclin & Tillmann, 2018). The meta-analysis suggested also that this advantage is not depending on higher general cognitive abilities of musicians. In fact, if this would be the explanation of the advantage of musicians, we should have observed a constant difference between groups in all the categories of stimuli presented, but this was not the case (e.g., visuospatial stimuli were associated to small or null effect sizes, whereas verbal stimuli were always associated to moderate effect sizes). On the other hand, we cannot exclude that musicians possess better strategies that can be applied only to tonal and verbal material, for example, but not to visuospatial stimuli. To conclude, STUDY 1 showed that there is an advantage of musicians over nonmusicians in memory tasks, especially in short-term and working memory, and that it varies depending on the type of stimuli that has to be remembered. This advantage seems not due to higher general cognitive skills, and not (or not only) to a publication bias.

STUDY 2 was driven by the interest of exploring short-term and working memory abilities, because several studies suggested that musicians had better verbal short-term and working memory than nonmusicians, and the fact that there seemed to be an advantage not only linked to music stimuli, but extended to other domains, such as the verbal one, was particularly fascinating (Bergman Nutley et al., 2014; Jakobson, Lewicky, & Kilgour, 2008). However, nobody ever disentangled if the advantage in verbal memory was driven by a general superior auditory memory. In fact, most of studies that presented verbal memory tasks, presented the items to be remembered auditorily. Because musicians possess fine auditory perception skills (e.g., Parbery-Clark et al., 2009; Spiegel & Watson, 1984), my first idea was to assess whether the advantage observed in verbal memory was still there when presenting verbal material visually. This idea was the base of STUDY 2, that aimed to investigate working memory of musicians and nonmusicians for verbal stimuli presented in three different modalities: visually, auditorily, and audiovisually. STUDY 2 revealed that musicians performed better than nonmusicians in the task, and the advantage (that was anyway small, half a digit span unit), was a bit more pronounced when the digits were presented auditorily and audiovisually. Moreover, there was still an advantage when the participants were asked to perform the articulatory suppression, that blocks the verbal rehearsal. Therefore, STUDY 2 showed a small advantage of musicians over nonmusicians in a digit span task, especially with digits presented auditorily and audiovisually, and this advantage seemed not linked to more efficient rehearsal strategies.

Given the results of the meta-analysis and of STUDY 2, the following step was to try to understand why musicians perform better than nonmusicians. One way to do this is trying to manipulate the type of music training. STUDY 3 focused on the type of training that musicians underwent. Specifically, the idea was to compare a formal training with a self-taught training. In fact, one of the most peculiar characteristics of the music training is the ability of read and convert music notation into movements and sounds. It is this multisensory integration one of the core characteristics of studying a music instrument. Nevertheless, there are musicians that learn to play without knowing how to read music notation, therefore they practice more the perceptual-motor aspect of the training. The idea was to compare musicians with this main difference: able to read and not able to read the music notation. This manipulation could help understanding whether it is only the formal training that is linked to the advantage observed in memory, or whether it is playing an instrument in general. STUDY 3 investigated three groups, reader musicians, nonreader musicians, and nonmusicians, who performed a short-term and working memory tasks, the digit span forward and backward. The tasks were again administered in three modalities, visually, auditorily, and audiovisually. STUDY 3 did not show an advantage of musicians over nonmusicians in any task. Nevertheless, an interaction between modality and group emerged in the forward span, and a significant small difference between groups was found between reader musicians and nonmusicians in the audiovisual presentation only. Qualitatively, it was possible to observe a tendency of both the two musicians' groups performing better than nonmusicians, especially with the auditory and audiovisual modality, and also in the backward span, but the difference was small (on average, less than one correct response), and not supported by the statistical analysis. This result was not expected, because literature, especially

STUDY 2, showed that the advantage in verbal memory tasks was moderate and quite consistent across different studies. Here, one possibility is that there were too many variables that attenuated the statistical power (i.e., the three modalities of presentation and the three groups). It is also possible that the numerosity of the sample, which was larger than the average of the studies presents in literature, revealed that the effect is not so strong as thought.

Finally, STUDY 4 aimed again to manipulate the type of stimuli. Most of studies in the literature investigated the tonal component of memory (in which, of course, musicians perform better than nonmusicians), the verbal component of memory (often with tasks presented auditorily), and the visuospatial component of memory. STUDY 4 investigated whether the advantage observed was linked to a general advantage in auditory memory, independently from the type of auditory stimuli presented, since the advantage seemed to be larger with auditory verbal stimuli and tonal stimuli than with visuospatial stimuli. The aim was therefore to compare visual and auditory stimuli, but including also nonverbal and nonmusical stimuli. Six different types of stimuli were created, to tap verbal memory (visual and auditory), and to tap nonverbal and nonmusical memory (visual and auditory). Moreover, the latter stimuli could be divided into two further categories: stimuli with contour and stimuli without contour. The contour manipulation was included to see whether musicians could generalize their skills of extracting automatically a contour (the up-down patterns) from melodies (Fujioka et al., 2004a) also with nonmusical material, both auditory and visual. The results showed that musicians performed better than nonmusicians only with the contour stimuli (in both modalities), and with the auditory nocontour stimuli. This suggested that musicians seem to be able to extract a contour structure also with nonmusical material, whereas nonmusicians are not able to do so. Moreover, it is likely that the fine perception skills of musicians allowed them to remember better other auditory nonmusical and nonverbal stimuli (i.e., the pink noise). Of course, the ability of extracting a contour relies also on the ability of discriminating the sounds. If the sounds appear all similar to each other, the person would find it impossible to hear a contour. But here musicians performed better also in the visual contour condition, therefore their fine auditory perception skills could not be the only explanation of their performance in the auditory contour condition. Surprisingly, here musicians did not perform better than nonmusicians in with the verbal stimuli, not even when the syllables were presented auditorily. This, again, is a case against the literature, and the reasons could be several: it could be possible that the task in this case was too easy for both groups, and since the difference is usually small, it did not emerge. It is also possible that the recognition paradigm was not often used in the past to test short-term memory of musicians, and maybe their advantage is more prominent in recall tasks. In any case, these results suggest that we are dealing with small effects that vary easily when changing variables and paradigms.

5.2 Theoretical implications

The present studies enlightened that the common distinction between verbal and visuospatial memory (for short-term and working memory) it is not sufficient. STUDY 1 suggested that there is a difference between groups mostly when the stimuli (i.e., the digits) are presented auditorily. STUDY 4 showed also that concerning nonverbal and nonmusical stimuli, musicians outperformed nonmusicians in the auditory task without contour, but no difference was found in the visual task without contour. A possibility is that there are different processes for auditory and visual memory, given that music expertise is linked to enhanced memory particularly for auditory sensory inputs, and that we cannot speak about memory basing only on the category of the content and ignoring the modality. Therefore, the present results would support more the memory models which are basing on the sensory input (Fougnie & Marois, 2011), with respect to the models that consider only the content to be remembered, such as the one of Baddeley (Baddeley & Hitch, 1974). Another possibility that the present results open - especially the results of STUDY 4 - is that it might be not important the modality of presentation of the task, neither the content of the task, but what matters could be the type of strategies that can be applied while performing the task. In STUDY 4 musicians performed better in the contour condition independently on the sensory modality of the task, suggesting that they could extract the pattern emerging from the relationship among the elements, and they could therefore remember the sequence as a structure, a sort of gestalt. It could be possible, therefore, that memory performance depends on what strategies can be used and when and less on the type of stimuli presented. It could be possible that the difference we observe between, for instance, auditory and visual memory, or between verbal and visuospatial memory, is not due to the fact that there are separate memory processes, but to the ability the individual develops in using optimal strategies for the specific task.

Concerning the debate about verbal and tonal memory, STUDY 2 showed that the effect sizes were of different dimensions depending on the type of stimuli (i.e., large for tonal stimuli; moderate for verbal stimuli), suggesting that the two materials are not overlapping in memory: if musicians can perform particularly well in tonal memory task, and not as much in verbal memory, this means that it is likely that the two materials are processed separately at some level.5.3

5.3 The contribute

The present dissertation could enlighten once again how expertise can be associated to enhanced abilities other than those that have been trained. Here, the meta-analysis, that was the first one that investigated memory skills of musicians, showed that musicians do perform better than nonmusicians in memory tasks, however, concerning verbal memory, it is not yet clear if the advantage is linked to a more general one of auditory memory, or if it extends to the verbal domain regardless the sensory input. In fact, here, two experimental studies (STUDY 2 and 3) did not show a large advantage of musicians over nonmusicians in short-term and working memory, but only small differences, especially when the task was presented auditorily; plus, STUDY 4 did not show any difference so far between groups in the verbal task. Nevertheless, the present works do suggest that a small advantage is there and veridical, and specifically it is present in musicians who had a formal training. This does not seem to be due to higher general cognitive abilities, or to possessing fine music perception skills. In fact, music aptitude was only partly related to the span performance, and mostly the relationship was between the subtests that involved clearly memory skills. Moreover, it is interesting to note that musicians do use their skills to perform better also with nonmusical material (STUDY 4), and that this happens automatically, suggesting that musicians learn implicit strategies for memorizing meaningful sequences of items. In fact, in the contour conditions of STUDY 4 musicians outperformed nonmusicians, and the difference between groups was large. Interestingly, they performed better in the visual contour than in the auditory one, showing how the ability to extract contour can improve the performance also with stimuli that are not usually associated with the feeling of contour (i.e., the luminance). This fact opens up new possibilities concerning the performance of musicians in the digit span task. Often, past studies found that musicians were performing better in digits recall tasks: a sequence of digit has a contour too, because every number is in a relationship (up -> bigger, down -> smaller) with each other. It could be interesting to investigate if musicians extract a contour also when remembering a sequence of digits. If so, this could explain the advantage observed in the past in tasks presenting digits.

5.4 Limitations

The present studies have some limitations. One is common to most of the other studies in the field, that is, the experimental studies here conducted are all quasi experiments. Even if we tried to control as many variables as possible, quasi experiments do not enable to investigate whether there is a causal relationship between music training and enhanced memory performance. In fact, musicians might be simply more motivated than nonmusicians and perform each task with more commitment. They can be also less likely to become tired after one hour of experimental session, that could lead to poorer performance. More control variables would have been needed, but here the choice of having, for example, only few subtests of the WAIS-IV, was driven mainly by time constrains. Whenever there are the resources, it is therefore important to run longitudinal studies, or, when running quasi experiments, it is important to include as many control variables as possible, and describe the participants and the method precisely, to allow replications. For example, one of the limits of quasi experiments that recruit musicians and nonmusicians is the following: is there a common definition of musicians? There is the need of an instrument that assesses the musicianship of musicians. Here the PROMS test could assess the general music aptitude of all participants, but it is not a test specifically developed for assessing the music abilities of musicians only (that of course are higher than those of nonmusicians). Moreover, other tools for investigating non-perceptual characteristics would be important, such as questionnaires to collect the details of the music training received, the method of study, the genre played (in jazz, for example, improvisation is a peculiar way of playing that require different skills than those for playing classical music), etc. Finally, larger sample sizes should be necessary to have a clear idea of the phenomenon that we are investigating, because usually they are small (e.g., around 20 participants per group).

5.6 Future steps

Plenty of work can be done. First of all, I believe that longitudinal studies, in which, for example, randomized groups of children are assigned to different training groups and observed during a long period of time, would be important to understand whether what we observe often in literature is a consequence of the music training or not. A longitudinal study could be designed also with adults, to see whether it is possible to produce improvements on cognitive tasks even if the person learns to play an instrument when the brain is fully developed. Secondly, if longitudinal experiments are not possible, another possibility is to run large quasi-experiments in which the years of music training are taken into account, to see whether there is a difference depending on the amount of training received (e.g., the performance increases with more years of training). Currently, we are designing a study with this aim in a multi-laboratory approach, to have a large number of participants and increase the reliability of the results. Moreover, a short-term music training could be developed to see whether it can have any immediate and delayed effect on the participants' cognitive skills. Concerning the last study, (STUDY 4), it would be interesting to design a second study to see whether the advantage of musicians over nonmusicians in the luminance task would disappear if teaching a specific strategy to nonmusicians, that is to associate a pitch to each level of luminance (as did musicians). If this will be the case, it would suggest that it is a matter of learning the best strategies, and not of having a superior memory per se.

5.7 Conclusion

The present studies investigated memory skills in musicians and nonmusicians, according to previous works that showed an advantage of musicians in memory task. The advantage observed in literature in verbal tasks and confirmed by the meta-analysis here included was only partially supported by the experimental studies' results. In fact, STUDY 2 found a small advantage mainly when the digits were presented auditorily, and STUDY 3 could not entirely replicate the findings, with a weaker advantage only in the audiovisual presentation. In the backward digit span, which investigates working memory, no difference between groups was found, in contrast with the results of the meta-analysis. In STUDY 4, no difference between groups in the verbal task was found, but a large difference with other auditory stimuli, nonverbal and nonmusical emerged. To sum up, the difference between musicians and nonmusicians here emerged are small and modulated by the modality of presentation of the stimuli. It could be possible that here there were too many conditions (e.g., the three different modalities of presentation of the digit span in STUDY 1 plus the three groups in STUDY 3) that hid partially the difference between groups. Or it is possible, that so far what was observed in literature is due to other variables, such as the modality of presentation, or personality traits (e.g., motivation). It is evident that details about this advantage are missing, and it would be useful to use different approaches when studying memory skills in musicians and nonmusicians. Basing, for example, only to Baddeley's working memory model would necessary limits the findings, if musicians possess, for example, enhanced auditory memory skills independently on the type of stimuli that have to be remembered. Whatever is the reason that musicians perform better than nonmusicians in memory tasks, it is important to acknowledge the potential that music training has, as a tool for exercising the cognitive functions.

REFERENCES

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? Psychological Bulletin, 131(1), 30–60. https://doi.org/10.1037/0033-2909.131.1.30

Aizenman, A. M., Gold, J. M., & Sekuler, R. (2017). Multisensory integration in short-term memory: Musicians do rock. Neuroscience, 1–11. https://doi.org/10.1016/j.neuroscience.2017.04.031

Albouy, P., Schulze, K., Caclin, A., & Tillmann, B. (2013). Does tonality boost short-term memory in congenital amusia? Brain Research, 1537, 224–232. https://doi.org/10.1016/j.brainres.2013.09.003

Alves, H., Voss, M., Boot, W. R., Deslandes, A., Cossich, V., Inacio Salles, J., & Kramer, A. F. (2013). Perceptual-cognitive expertise in elite volleyball players. Frontiers in psychology, 4, 36.

Amer, T., Kalender, B., Hasher, L., Trehub, S. E., & Wong, Y. (2013). Do Older Professional Musicians Have Cognitive Advantages? PLoS ONE, 8(8), 1–8. https://doi.org/10.1371/journal.pone.0071630

Anaya, E. M., Pisoni, D. B., & Kronenberger, W. G. (2016). Long-term musical experience and auditory and visual perceptual abilities under adverse conditions. The Journal of the Acoustical Society of America, 140(3), 2074–2081. https://doi.org/10.1121/1.4962628

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. Journal of Experimental Child Psychology. https://doi.org/10.1016/S0022-0965(02)00124-8

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience*, *4*(10), 829.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Sciences, 4(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. (1974). Working memory. The Psychology of Learning and Motivation Advances in Research and Theory. https://doi.org/10.1016/S0079-7421(08)60452-1

Barwick, J., Valentine, E., West, R., & Wilding, J. (1989). relations between reading and musical abilities. British Journal of Educational Psychology. https://doi.org/10.1111/j.2044-8279.1989.tb03097.x

Bergman Nutley, S., Darki, F., & Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. Frontiers in Human Neuroscience, 7(1), 1–9. https://doi.org/10.3389/fnhum.2013.00926

Bialystok, E., & DePape, A. M. (2009). Musical Expertise, Bilingualism, and Executive Functioning. Journal of Experimental Psychology: Human Perception and Performance, 35(2), 565–574. https://doi.org/10.1037/a0012735

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone Language Speakers and Musicians Share Enhanced Perceptual and Cognitive Abilities for Musical Pitch: Evidence for Bidirectionality between the Domains of Language and Music. PLoS ONE, 8(4). https://doi.org/10.1371/journal.pone.0060676

Blachman, B. A. (2000). Phonological awareness. *Handbook of reading research*, *3*, 483-502.

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. The Journal of the Acoustical Society of America, 137(1), 378–387. https://doi.org/10.1121/1.4904537

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. The handbook of research synthesis and meta-analysis, 2, 221-235.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to Meta-Analysis. John Wiley & Sons. https://doi.org/10.1002/9780470743386

Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. Psychology of music, 31(2), 123-138.

Caclin, A., & Tillmann, B. (2018). Musical and verbal short-term memory: Insights from neurodevelopmental and neurological disorders. Annals of the New York Academy of Sciences. https://doi.org/10.1111/nyas.13733

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2004). Action observation and acquired motor skills: an FMRI study with expert dancers. Cerebral cortex, 15(8), 1243-1249.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., ... & Altmejd, A. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour, 1.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology. https://doi.org/10.1016/0022-0965(82)90054-6

Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. Nature, 396(6707), 128. https://doi.org/10.1038/24075

Chang, Y. (2014). Reorganization and plastic changes of the human brain associated with skill learning and expertise. Frontiers in human neuroscience, 8, 35.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive psychology, 4(1), 55-81.

Chi, M. T., Glaser, R., & Farr, M. J. (2014). The nature of expertise. Psychology Press.

Chin, T., & Rickard, N. S. (2012). The Music USE (MUSE) Questionnaire: An Instrument to Measure Engagement in Music. Music Perception: An Interdisciplinary Journal. https://doi.org/10.1525/mp.2012.29.4.429

Christoffels, I. K., De Groot, A. M., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. Journal of Memory and Language, 54(3), 324-345.

Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. PLoS ONE, 11(7), 1–17. https://doi.org/10.1371/journal.pone.0157638

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: L. Lawrence Earlbaum Associates, 2. . https://doi.org/10.1234/12345678

Cohen, M. A., Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). Auditory and visual memory in musicians and nonmusicians. Psychonomic Bulletin and Review, 18(3), 586–591. https://doi.org/10.3758/s13423-011-0074-0

Corrigall, K. A., & Trainor, L. J. (2011). Associations between length of music training and reading skills in children. Music Perception: An Interdisciplinary Journal, 29(2), 147-155.

Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. Journal of research in music education, 47(3), 198-212.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, *169*, 323-338. https://doi.org/10.1016/S0079-6123(07)00020-9

Crottaz-Herbette, S., Anagnoson, R. T., & Menon, V. (2004). Modality effects in verbal working memory: Differential prefrontal and parietal responses to auditory and visual stimuli. NeuroImage, 21(1), 340–351. https://doi.org/10.1016/j.neuroimage.2003.09.019

De Groot, A. D. (2014). Thought and choice in chess (Vol. 4). Walter de Gruyter GmbH & Co KG.

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: A relation mediated by executive functions. Music Perception: An Interdisciplinary Journal, 29(2), 195-201.

Del Re AC, Hoyt WT. MAd: meta-analysis with mean differences. R package version 0.9 ed. 2010. Available from: https://CRAN.R-project.org/package=MAd

Deutsch, D. (2013). The Psychology of Music (Third edition). Elsevierhttps://doi.org/10.1016/C2009-0-62532-0

Di Paola, M., Caltagirone, C., & Petrosini, L. (2013). Prolonged rock climbing activity induces structural changes in cerebellum and parietal lobe. Human brain mapping, 34(10), 2707-2714.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: changes in grey matter induced by training. Nature, 427(6972), 311.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. Perception & Psychophysics, 53(1), 25-33.

Duval, S. (2006). The Trim and Fill Method. In Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. https://doi.org/10.1002/0470870168.ch8

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. Science. https://doi.org/10.1126/science.270.5234.305

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: a general overview. Academic emergency medicine, 15(11), 988-994.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. American Psychologist. https://doi.org/10.1037/0003-066X.49.8.725

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. Psychological Review. https://doi.org/10.1037/0033-295X.102.2.211

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological review, 100(3), 363.

Fougnie, D., & Marois, R. (2011). What Limits Working Memory Capacity? Evidence for Modality-Specific Sources to the Simultaneous Storage of Visual and Auditory Arrays. Journal of Experimental Psychology: Learning Memory and Cognition, 37(6), 1329–1341. https://doi.org/10.1037/a0024834

Franěk, M., Mates, J., Radil, T., Beck, K., & Pöppel, E. (1991). Finger tapping in musicians and nonmusicians. International Journal of psychophysiology, 11(3), 277-279.

Franklin, M. S., Moore, K. S., Yip, C. Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. Psychology of Music, 36(3), 353–365. https://doi.org/10.1177/0305735607086044

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. Journal of Cognitive Neuroscience. https://doi.org/10.1162/0898929041502706

Furley, P., & Memmert, D. (2010). Differences in spatial working memory as a function of team sports expertise: the Corsi Block-tapping task in sport psychological assessment. Perceptual and motor skills, 110(3), 801-808.

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. Neuropsychologia, 49(5), 1083–1094. https://doi.org/10.1016/j.neuropsychologia.2011.02.001

Gibson, C., Folley, B. S., & Park, S. (2009). Enhanced divergent thinking and creativity in musicians: A behavioral and near-infrared spectroscopy study. Brain and cognition, 69(1), 162-169.

Giofrè, D., Mammarella, I. C., & Cornoldi, C. (2013). The structure of working memory and how it relates to intelligence in children. Intelligence, 41(5), 396–406. https://doi.org/10.1016/j.intell.2013.06.006

Gold, J. M., Aizenman, A., Bond, S. M., & Sekuler, R. (2014). Memory and incidental learning for visual frozen noise sequences. Vision Research, 99, 19–36. https://doi.org/10.1016/j.visres.2013.09.005

Gorin, S., Kowialiewski, B., & Majerus, S. (2016). Domain-generality of timing-based serial order processes in short-term memory: New insights from musical and verbal domains. PLoS ONE, 11(12), 1–25. https://doi.org/10.1371/journal.pone.0168699

Goswami, U., Huss, M., Mead, N., Fosker, T., & Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: Significant longitudinal relations with word reading and reading comprehension. Cortex. https://doi.org/10.1016/j.cortex.2012.05.005

Grassi, M., & Borella, E. (2013). The role of auditory abilities in basic mechanisms of cognition in older adults. Frontiers in Aging Neuroscience. https://doi.org/10.3389/fnagi.2013.00059

Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. Journal of Research in Music Education, 53(3), 199-209.

Habib, M., & Besson, M. (2009). What do Music Training and Musical Experience Teach Us About Brain Plasticity? Music Perception: An Interdisciplinary Journal. https://doi.org/10.1525/mp.2009.26.3.279

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(July), 1–11. https://doi.org/10.3389/fnhum.2012.00198

Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. Psychology of Music, 41(6), 779–793. https://doi.org/10.1177/0305735612452186

Helmbold, N., Rammsayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. Journal of Individual Differences, 26(2), 74-85.

Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. Neuron. https://doi.org/10.1016/j.neuron.2012.10.011

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. BMJ : British Medical Journal. https://doi.org/10.1136/bmj.327.7414.557

Hill, N. M., & Schneider, W. (2006). Brain changes in the development of expertise: Neuroanatomical and neurophysiological evidence about skill-based adaptations. The Cambridge handbook of expertise and expert performance, 653-682.

Hirel, C., Nighoghossian, N., Lévêque, Y., Hannoun, S., Fornoni, L., Daligault, S., … Caclin, A. (2017). Verbal and musical short-term memory: Variety of auditory disorders after stroke. Brain and Cognition, 113, 10–22. https://doi.org/10.1016/j.bandc.2017.01.003

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. Neuropsychology, 17(3), 439–450. https://doi.org/10.1037/0894-4105.17.3.439

Huang, Z., Zhang, J. X., Yang, Z., Dong, G., Wu, J., Chan, A. S., & Weng, X. (2010). Verbal memory retrieval engages visual cortex in musicians. Neuroscience, 168(1), 179–189. https://doi.org/10.1016/j.neuroscience.2010.03.027

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). The effects of musical training on structural brain development: A longitudinal study. In Annals of the New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.2009.04852.x

Icano, A. O. P., Iani, C., RubichI, S., Ricciardelli, P., BorghI, A. M., & Nicoletti, R. (2010). Real-life motor training modifies spatial performance: The advantage of being drummers. American Journal of Psychology. https://doi.org/10.5406/amerjpsyc.123.2.0169

Jacini, W. F., Cannonieri, G. C., Fernandes, P. T., Bonilha, L., Cendes, F., & Li, L. M. (2009). Can exercise shape your brain? Cortical differences associated with judo practice. Journal of Science and Medicine in Sport, 12(6), 688-690.

Jakobson L. S., Lewycky S. T., Kilgour A. R., & Stoesz, B. M. (2008). Memory for verbal and visual material in highly trained musicians. Music Perception, 26(1), 41–56. Retrieved from https://doi.org/10.1525/mp.2008.26.1.41

Jäncke, L. (2009). The plastic human brain. Restorative Neurology and Neuroscience. https://doi.org/10.3233/RNN-2009-0519

Jäncke, L., Koeneke, S., Hoppe, A., Rominger, C., & Hänggi, J. (2009). The architecture of the golfer's brain. PloS one, 4(3), e4785.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? Journal of basic and clinical physiology and pharmacology, 12(2), 125-144.

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. Journal of Experimental Psychology: General. https://doi.org/10.1037/0096-3445.133.3.355

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. Nature reviews neuroscience, 11(8), 599.

Lee, Y. shiow, Lu, M. ju, & Ko, H. ping. (2007). Effects of skill training on working memory capacity. Learning and Instruction, 17(3), 336–344. https://doi.org/10.1016/j.learninstruc.2007.02.010

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.070039597

Majerus, S., & Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: The importance of serial order. Frontiers in Psychology, 7(OCT). https://doi.org/10.3389/fpsyg.2016.01522

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. Journal of Experimental Child Psychology, 93(2), 95–119. https://doi.org/10.1016/j.jecp.2005.07.005

Mandikal Vasuki, P. R., Sharma, M., Demuth, K., & Arciuli, J. (2016). Musicians’ edge: A comparison of auditory processing, cognitive abilities and statistical learning. Hearing Research, 342, 112–123. https://doi.org/10.1016/j.heares.2016.10.008

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. Journal of Cognitive Neuroscience, 23(10), 2701-2715.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. British Journal of Psychology. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Marks, L. E. (1987). On Cross-Modal Similarity: Auditory-Visual Interactions in Speeded Discrimination. Journal of Experimental Psychology: Human Perception and Performance. https://doi.org/10.1037/0096-1523.13.3.384

Marques, C., Moreno, S., Luís Castro, S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. Journal of Cognitive Neuroscience, 19(9), 1453-1463.

Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. Journal of Experimental Child Psychology, 111(4), 708–723. https://doi.org/10.1016/j.jecp.2011.11.007

Martinez Perez, T., Majerus, S., & Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. Research in Developmental Disabilities. https://doi.org/10.1016/j.ridd.2013.04.005

Martínez-Montes, E., Hernández-Pérez, H., Chobert, J., Morgado-Rodríguez, L., Suárez-Murias, C., Valdés-Sosa, P. A., & Besson, M. (2013). Musical expertise and foreign speech perception. Frontiers in systems neuroscience, 7, 84.

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. Current Directions in Psychological Science. https://doi.org/10.1111/1467-8721.00116

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? Psychological Science. https://doi.org/10.1111/j.1467-9280.2008.02235.x

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. Hearing Research. https://doi.org/10.1016/j.heares.2006.05.004

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. Brain Research. https://doi.org/10.1016/j.brainres.2007.11.042

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

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine. https://doi.org/10.1371/journal.pmed.1000097

Monahan, C. B., Kendall, R. A., & Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. Perception & Psychophysics, 41(6), 576–600. https://doi.org/10.3758/BF03210491

Morrongiello, B. A., & Roes, C. L. (1990). Developmental changes in children's perception of musical sequences: Effects of musical training. Developmental Psychology, 26(5), 814.

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. https://doi.org/10.1371/journal.pone.0089642

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. Nature Human Behaviour. https://doi.org/10.1038/s41562-016-0021

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. Nature Reviews Neuroscience, 3(6), 473–478. https://doi.org/10.1038/nrn843

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., & Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? Brain and Cognition. https://doi.org/10.1016/j.bandc.2005.05.009

Okhrei, A., Kutsenko, T., & Makarchuk, M. (2017). Performance of working memory of musicians and non-musicians in tests with letters, digits, and geometrical shapes. Biologija, 62(4), 207–215. https://doi.org/10.6001/biologija.v62i4.3408

Overy, K., Nicolson, R. I., Fawcett, A. J., & Clarke, E. F. (2003). Dyslexia and music: Measuring musical timing skills. Dyslexia. https://doi.org/10.1002/dys.233

Pačesová, P., Šmela, P., Kraček, S., Kukurová, K., & Plevková, L. (2018). Cognitive function of young male tennis players and non-athletes.

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. PLoS ONE, 5(6). https://doi.org/10.1371/journal.pone.0011120

Paraskevopoulos, E., Kraneburg, A., Herholz, S. C., Bamidis, P. D., & Pantev, C. (2015). Musical expertise is related to altered functional connectivity during audiovisual integration. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1510662112

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Musical Expertise Induces Audiovisual Integration of Abstract Congruency Rules. Journal of Neuroscience. https://doi.org/10.1523/JNEUROSCI.1947-12.2012

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-In-noise. Ear and Hearing, 30(6), 653–661. https://doi.org/10.1097/AUD.0b013e3181b412e9

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. PLoS ONE, 6(5). https://doi.org/10.1371/journal.pone.0018082

Pascual-Leone, A., & Torres, F. (1993). Plasticity of the sensorimotor cortex representation of the reading finger in Braille readers. Brain, 116(1), 39-52.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. Annu. Rev. Neurosci., 28, 377-401.

Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., & Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. Journal of neurophysiology, 74(3), 1037-1045.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-Term Memory, Working Memory, and Inhibitory Control in Children with Difficulties in Arithmetic Problem Solving. Journal of Experimental Child Psychology, 80(1), 44–57. https://doi.org/10.1006/jecp.2000.2626

Patel, A. D. (2003). Language, music, syntax and the brain. Nature Neuroscience. https://doi.org/10.1038/nn1082

Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. Trends in Cognitive Sciences. https://doi.org/10.1016/j.tics.2007.08.003

Pearce, A. J., Thickbroom, G. W., Byrnes, M. L., & Mastaglia, F. L. (2000). Functional reorganisation of the corticomotor projection to the hand in skilled racquet players. Experimental Brain Research, 130(2), 238-243.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. Nature neuroscience, 6(7), 688.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital Amusia: A disorder of fine-grained pitch discrimination. Neuron. https://doi.org/10.1016/S0896-6273(01)00580-3

Peretz, I., Cummings, S., & Dubé, M.-P. (2007). The Genetics of Congenital Amusia (Tone Deafness): A Family-Aggregation Study. The American Journal of Human Genetics, 81(3), 582–588. https://doi.org/10.1086/521337

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. Brain, 117(6), 1283-1301.

Perez, T. M., Majerus, S., Mahot, A., & Poncelet, M. (2012). Evidence for a specific impairment of serial order short-term memory in dyslexic children. Dyslexia, 18(2), 94–109. https://doi.org/10.1002/dys.1438

Peynirciogˇlu, Z. F., Durgunogˇlu, A. Y., & Úney-Küsefogˇlu, B. (2002). Phonological awareness and musical aptitude. Journal of Research in Reading. https://doi.org/10.1111/1467-9817.00159

Piro, J. M., & Ortiz, C. (2009). The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. Psychology of Music, 37(3), 325-347.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). Cross-modal melodic contour similarity. Canadian Acoustics.

R Development Core Team, R. (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7

Ramachandra, V., Meighan, C., & Gradzki, J. (2012). The Impact of Musical Training on the Phonological Memory and the Central Executive: A Brief Report. North American Journal of Psychology, 14(3), 541–548.

Rammsayer, T., & Altenm�ller, E. (2006). Temporal Information Processing in Musicians and Nonmusicians. Music Perception. https://doi.org/10.1525/mp.2006.24.1.37

Rauscher, F. H. (2002). Mozart and the mind: Factual and fictional effects of musical enrichment. In J. Aronson (Ed.), Improving academic achievement: Impact of psychological factors on education, (pp. 267-278). New York, NY: Academic Press.

Rauscher, F. H., & Zupan, M. A. (2000). Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: A field experiment. Early Childhood Research Quarterly, 15(2), 215-228.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. Psychological Science, 12(1), 48-55.

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. Psychology of Music, 42(2), 284–298. https://doi.org/10.1177/0305735612471239

Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. Frontiers in Psychology, 3(DEC), 1–9. https://doi.org/10.3389/fpsyg.2012.00572

Rodrigues, A. C., Loureiro, M. A., & Caramelli, P. (2013). Long-term musical training may improve different forms of visual attention ability. Brain and cognition, 82(3), 229-235.

Rodrigues, A. C., Loureiro, M., & Caramelli, P. (2014). Visual memory in musicians and non-musicians. Frontiers in Human Neuroscience, 8(June), 1–10. https://doi.org/10.3389/fnhum.2014.00424

Roever C. bayesmeta: Bayesian Random-Effects Meta-Analysis. 2015. Available from: https://CRAN. R-project.org/package=bayesmeta

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. The Urban Review. https://doi.org/10.1007/BF02322211

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. https://doi.org/10.1002/0470870168

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. PloS one, 9(1), e86980.

Schellenberg, E. G. (2004). Music lessons enhance IQ. Psychological science, 15(8), 511-514.

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. Journal of Educational Psychology, 98(2), 457.

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. British Journal of Psychology, 102(3), 283-302.

Schellenberg, E. G., & Mankarious, M. (2012). Music training and emotion comprehension in childhood. Emotion, 12(5), 887.

Schellenberg, E. G., & Peretz, I. (2008). Music, language and cognition: unresolved issues. Trends in cognitive sciences, 12(2), 45-46.

Schellenberg, E. G., & W. Weiss, M. (2013). 12 - Music and Cognitive Abilities. The Psychology of Music (Third Edition). https://doi.org/10.1016/B978-0-12-381460-9.00012-2

Schiavio, A., & Timmers, R. (2016). Motor and audiovisual learning consolidate auditory memory of tonally ambiguous melodies. Music Perception. https://doi.org/10.1525/mp.2016.34.1.21

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. Neuropsychologia. https://doi.org/10.1016/0028-3932(95)00045-5

Scholz, J., Klein, M. C., Behrens, T. E., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. Nature neuroscience, 12(11), 1370.

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J. L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. Neuroimage, 51(1), 450-461.

Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. Music Perception: An Interdisciplinary Journal, 29(3), 255-267.

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. European Journal of Neuroscience, 33(1), 189–196. https://doi.org/10.1111/j.1460-9568.2010.07470.x

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Human Brain Mapping, 32(5), 771–783. https://doi.org/10.1002/hbm.21060

Seashore, C. E. (1915). The measurement of musical talent. The Musical Quarterly, *1*(1), 129-148.

Seashore, C. E., Lewis, D., & Saetveit, J. G. (1960). Measures of Musical Talents (test). Psychological Corporation.

Sekiguchi, A., Yokoyama, S., Kasahara, S., Yomogida, Y., Takeuchi, H., Ogawa, T., ... & Kawashima, R. (2011). Neural bases of a specific strategy for visuospatial processing in rugby players. Medicine and science in sports and exercise, 43(10), 1857-1862.

Slagter, H. A., Davidson, R. J., & Lutz, A. (2011). Mental training as a tool in the neuroscientific study of brain and cognitive plasticity. Frontiers in human neuroscience, 5, 17.

Slevc, L. R., & Miyake, A. (2006). Individual Differences in Second- Language Proficiency. Psychological Science. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. Neuroimage, 17(3), 1613-1622.

Sluming, V., Brooks, J., Howard, M., Downes, J. J., & Roberts, N. (2007). Broca's area supports enhanced visuospatial cognition in orchestral musicians. Journal of Neuroscience, 27(14), 3799-3806.

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. The Journal of the Acoustical Society of America. https://doi.org/10.1121/1.391605

Stoesz, B. M., Jakobson, L. S., Kilgour, A. R., & Lewycky, S. T. (2007). Local processing advantage in musicians: Evidence from disembedding and constructional tasks. Music Perception: An Interdisciplinary Journal, 25(2), 153-165.

Strait, D. L., & Kraus, N. (2014). Biological impact of auditory expertise across the life span: Musicians as a model of auditory learning. Hearing Research. https://doi.org/10.1016/j.heares.2013.08.004

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. Brain and language, 123(3), 191-201.

Suárez, L., Elangovan, S., & Au, A. (2016). Cross-sectional study on the relationship between music training and working memory in adults. Australian Journal of Psychology, 68(1), 38–46. https://doi.org/10.1111/ajpy.12087

Suresh, K., & Chandrashekara, S. (2012). Sample size estimation and power analysis for clinical research studies. Journal of Human Reproductive Sciences. https://doi.org/10.4103/0974-1208.97779

Talamini, F., Carretti, B., & Grassi, M. (2016). The working memory of musicians and nonmusicians. Music Perception, 34(2). https://doi.org/10.1525/MP.2016.34.2.183

Talamini, F., Grassi, M., Toffalini, E., Santoni, R., & Carretti, B. (2018). Learning a second language: Can music aptitude or music training have a role? Learning and Individual Differences, 64(January 2017), 1–7. https://doi.org/10.1016/j.lindif.2018.04.003

Tang, Y. Y., Lu, Q., Fan, M., Yang, Y., & Posner, M. I. (2012). Mechanisms of white matter changes induced by meditation. Proceedings of the National Academy of Sciences, 109(26), 10570-10574.

Taylor, A. C., & Dewhurst, S. A. (2017). Investigating the influence of music training on verbal memory. Psychology of Music, 45(6), 814–820. https://doi.org/10.1177/0305735617690246

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. Experimental Brain Research. https://doi.org/10.1007/s00221-004-2044-5

Tierney, A. T., Bergeson-Dana, T. R., & Pisoni, D. B. (2009). Effects of Early Musical Experience on Auditory Sequence Memory. Empirical Musicology Review, 6(4), 247–253. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial

Tillmann, B., Lévêque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. Brain Research. https://doi.org/10.1016/j.brainres.2015.10.035

Tillmann, B., Schulze, K., & Foxton, J. M. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. Brain and Cognition, 71(3), 259–264. https://doi.org/10.1016/j.bandc.2009.08.003

Travis, F., Harung, H. S., & Lagrosen, Y. (2011). Moral development, executive functioning, peak experiences and brain patterns in professional and amateur classical musicians: Interpreted in light of a Unified Theory of Performance. Consciousness and Cognition. https://doi.org/10.1016/j.concog.2011.03.020

Verburgh L, Scherder EJA, van Lange PA, Oosterlaan J (2014) Executive Functioning in Highly Talented Soccer Players. PLoS ONE 9(3): e91254. https://doi.org/10.1371/journal.pone.0091254

Viechtbauer W, Viechtbauer MW. Package `metafor'. The Comprehensive R Archive Network. Package `metafor'. 2015. Available from: http://cran.r-project.org/web/packages/metafor/metafor.pdf.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. https://doi.org/10.1103/PhysRevB.91.121108

Voss, M. W., Kramer, A. F., Basak, C., Prakash, R. S., & Roberts, B. (2010). Are expert athletes 'expert'in the cognitive laboratory? A meta-analytic review of cognition and sport expertise. Applied Cognitive Psychology, 24(6), 812-826.

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. https://doi.org/10.1016/j.lindif.2010.02.004

Wechsler, D. (2003). Wechsler intelligence scale for children–Fourth Edition (WISC-IV). *San Antonio, TX: The Psychological Corporation*.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). San Antonio.

Weiss, A. H., Biron, T., Lieder, I., Granot, R. Y., & Ahissar, M. (2014). Spatial vision is superior in musicians when memory plays a role. Journal of Vision, 14(9), 18–18. https://doi.org/10.1167/14.9.18

Williams, A. M. (2000). Perceptual skill in soccer: Implications for talent identification and development. Journal of sports sciences, 18(9), 737-750.

Williams, M., & Davids, K. (1995). Declarative knowledge in sport: A by-product of experience or a characteristic of expertise?. Journal of sport and exercise psychology, 17(3), 259-275.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. Memory and Cognition, 38(2), 163–175. https://doi.org/10.3758/MC.38.2.163

Yang, H., Lu, J., Gong, D., & Yao, D. (2016). How do musical tonality and experience affect visual working memory? NeuroReport, 27(2), 94–98. https://doi.org/10.1097/WNR.0000000000000503

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. PLoS ONE, 9(6). https://doi.org/10.1371/journal.pone.0099868