Music and Cognitive Abilities

E. Glenn Schellenberg and Michael W. Weiss
Department of Psychology, University of Toronto, Mississauga, Ontario, Canada

I. Introduction

In this chapter, we review the available evidence concerning associations between music and cognitive abilities. We use the term “cognitive abilities” to refer to all aspects of cognition (e.g., memory, language, visuospatial abilities) including general intelligence. We use the term “music” as an all-encompassing one that includes music aptitude, music listening, and music lessons, and we use the term “associations” because it does not imply causation. Our focus is on documented associations—regardless of the direction of causation—between cognitive abilities, on the one hand, and music aptitude, music listening, and music lessons, on the other. In each case, we examine the possibility of a causal relationship between music and cognition.

The chapter is divided into four main sections: music aptitude and cognitive abilities, cognitive abilities after music listening (the so-called Mozart effect), background music and cognitive abilities (i.e., cognitive abilities while listening to music), and music training and cognitive abilities (i.e., cognitive abilities as a function of music training). Our review focuses on articles published in English with behavioral outcome measures. Links between music and brain function or structure are discussed in Chapters 13 and 14 (this volume).

II. Music Aptitude and Cognitive Abilities

Music aptitude refers to natural music abilities or the innate potential to succeed as a musician. One school of thought (e.g., Ericsson, Krampe, & Tesch-Römer, 1993; Howe, Davidson, & Sloboda, 1998) contends that innate music talent (i.e., aptitude plus a demonstrated ability to perform music) does not account for variations in levels of musicality. Rather, expert levels can be achieved by anyone who starts early enough and works hard enough. In short, practice makes perfect (cf. Meinz & Hambrick, 2010). The debate about the existence of music talent or aptitude is beyond the scope of the present chapter. We assume that music aptitude exists, that it varies among individuals, and that aptitude is something that tests of music
aptitude measure. Although this definition is circular, our principal focus is on whether tests of music aptitude measure an ability that is independent of or associated with other cognitive abilities.

The issue of music aptitude as an ability distinct from other cognitive abilities is closely related to concepts of modularity (Fodor, 1983; Peretz, 2009; Peretz & Coltheart, 2003) and multiple intelligences (Gardner, 1983, 1999). The notion of modularity proposes that (1) the brain has specialized modules for processing different kinds of information, (2) domain-specific information (re: language, faces, music, and so on) is processed automatically by the appropriate module, and (3) each module functions independently of other modules (Fodor, 1983). Gardner (1983, 1999) posits similarly that intelligence is a multidimensional construct. In the original formulation of his theory of multiple intelligences, he specified seven distinct intelligences: bodily-kinesthetic, interpersonal, intrapersonal, linguistic, logical-mathematical, spatial, and most importantly, musical intelligence. From either the modularity or multiple intelligences perspective, music aptitude should be distinct from other abilities.

The typical task on tests of music aptitude involves presenting two short melodies (or two short rhythms) on each trial. Listeners are asked whether the second melody (or rhythm) is the same as or different from the first. After several trials, a score is calculated separately for each test. An aggregate score can also be calculated by averaging across tests. The origin of music-aptitude testing is often attributed to Seashore (1919, 1960). Seashore’s test has six subtests, including pitch and rhythm tasks as well as subtests of loudness, meter, timbre, tonal memory, and an aggregate measure of general music aptitude. In North America, one of the most well-known measures is Gordon’s (1965) Music Aptitude Profile (MAP), which has seven subtests. Gordon later simplified his approach, forming three different tests based on grade level: the Primary Measures of Music Audiation (PMMA, kindergarten to third grade; Gordon, 1979), the Intermediate Measures of Music Audiation (IMMA, first to sixth grade; Gordon, 1982), and the Advanced Measures of Music Audiation (AMMA, seventh grade and higher; Gordon, 1989). Each test has only two subtests (pitch and rhythm). In the United Kingdom, Wing’s (1962) Musical Aptitude Test has been used frequently, as has Bentley’s (1966) Measures of Musical Abilities. All of the aptitude tests and their corresponding subtests tend to be correlated at moderate to high levels (e.g., Gordon, 1969; McLeish, 1968; Vispoel, 1992; Young, 1972, 1973). Unfortunately, there is no test of music aptitude that is considered to be the “gold standard.” Consequently, the particular test varies from study to study, which undoubtedly contributes to inconsistent findings.

In tests of the validity of aptitude measures, criterion variables (i.e., those that should be correlated with aptitude) also vary across studies. Often, construct validity is tested by examining associations between aptitude scores and a teacher’s or parent’s subjective rating of the participant’s “musicality” (Davies, 1971; Drake, 1954; Harrington, 1969; Tarrell, 1965; Young, 1972, 1976). These correlations are typically positive but small to moderate in size, which is not surprising because aptitude tests present participants with short auditory sequences.
that are a far cry from actual musical pieces. Music training is seldom used as a criterion variable precisely because aptitude tests are supposed to measure natural music ability independent of training. Nevertheless, when musically trained and untrained participants are compared, the trained group typically has higher aptitude scores (Bentley, 1970; Davies, 1971; Drake, 1954; Flohr, 1981; Forgeard, Winner, Norton, & Schlaug, 2008; Gordon, 1969, 1980, 2001; Hassler, 1992; Hassler, Birbaumer, & Feil, 1985; Isaacs & Trofimovich, 2011; Milovanov, Pietilä, Tervaniemi, & Esquef, 2010; Milovanov, Tervaniemi, Takio, & Hämäläinen, 2007; Posedel, Emery, Souza, & Fountain, 2011; Tsang & Conrad, 2011; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010).

The major question discussed here is whether music aptitude is related to nonmusical cognitive abilities. The relevant studies have examined associations between music aptitude and language skills, mathematical abilities, and general intelligence. Links between music aptitude and nonmusical abilities are problematic for theorists who posit modularity for music (Peretz, 2009; Peretz & Coltheart, 2003) or that music ability represents an intelligence distinct from other abilities (Gardner, 1983, 1999).

### A. Music Aptitude and Language

The skills required to perform well on music-aptitude tests may be useful for other auditory tasks, such as those involving language. For example, understanding spoken language requires the listener to perceive the relevant segmental units that make up words, including the sounds associated with particular consonants and vowels. One example of a test of phonological (or phonemic) awareness asks participants to repeat a word with the initial or final phoneme or syllable deleted: spent with its first phoneme deleted is pent, cold with its final phoneme deleted is coal, and rainbow with its second syllable deleted is rain. Children’s music aptitude is positively associated with performance on these sound-deletion tasks (Forgeard, Schlaug, et al., 2008; Huss, Verney, Fosker, Mead, & Goswami, 2011; Norton et al., 2005; Peynircioğlu, Durgunoğlu, & Öney-Küsefoğlu, 2002). Adults who have poor pitch perception also exhibit deficits in phonological processing (Jones, Lucker, Zalewski, Brewer, & Drayna, 2009).

Because phonological awareness is predictive of reading ability (Bradley & Bryant, 1983; Stahl & Murray, 1994), researchers have examined whether music aptitude is not only associated with phonological awareness but also with actual measures of reading. In one study of 4- and 5-year-olds, phonological awareness was associated with music aptitude as well as with reading ability (Anvari, Trainor, Woodside, & Levy, 2002). Music aptitude and reading ability were correlated among 4-year-olds. Among 5-year-olds, however, performance on the pitch part of aptitude tests was correlated with reading ability but rhythmic aptitude was not. More importantly, for both age groups, the association between aptitude and reading was evident even when phonological awareness was held constant, which points to a direct link between music aptitude and reading that is independent of phonological awareness. This link disappeared among 4-year-olds when working
memory or vocabulary was controlled in addition to phonological awareness, but it remained evident when arithmetic ability was held constant. For 5-year-olds, the link between pitch-based aptitude and reading was still significant when phonological awareness and working memory, vocabulary, or arithmetical ability was controlled, but the researchers did not test whether the association was evident when all of the possible confounding variables were held constant simultaneously. Moreover, because there was no measure of general intelligence, the association between aptitude and reading may have stemmed from the fact that high-ability children tend to perform well on tests of music aptitude and reading.

In another study of 5-year-olds, performance on tests of pitch-based aptitude, phonological awareness, and reading abilities was intercorrelated (Lamb & Gregory, 1993). Correlations between phonological awareness and aptitude remained evident when intelligence was held constant. In a second study of 7- to 10-year-olds, reading ability was correlated with performance on one of the Bentley subtests (chord analysis—judging how many notes are presented simultaneously) even when age and IQ were controlled (Barwick, Valentine, West, & Wilding, 1989). In a third study of 8- to 11-year-olds, performance on two of the Bentley subtests (chord analysis, pitch-based aptitude) was predictive of reading ability with age and IQ held constant (Barwick et al., 1989). In the latter two studies, performance on a test of rhythmic aptitude was not associated with reading ability.

When Douglas and Willatts (1994) tested 7- and 8-year-olds, positive correlations were observed among tests of pitch and rhythmic aptitude, vocabulary, reading, and spelling. When vocabulary abilities were held constant, the association between rhythm and reading abilities remained significant, as did the correlation between rhythm and spelling abilities. By contrast, correlations between pitch aptitude and reading or spelling disappeared when vocabulary was controlled. Thus, rhythmic aptitude rather than pitch aptitude was more strongly associated with reading and spelling abilities, a finding that conflicts with others (Anvari et al., 2002; Barwick et al., 1989; Lamb & Gregory, 1993). Studies of children with dyslexia also suggest that deficits in rhythmic aptitude predict problems with phonological awareness and reading (Huss et al., 2011; Overy, 2000, 2003; Overy, Nicolson, Fawcett, & Clarke, 2003). The issue is further complicated by a report indicating that pitch but not rhythmic aptitude predicts reading ability for untrained but not for musically trained children (Tsang & Conrad, 2011).

Individual differences in music aptitude are associated with the ability to acquire a second language (e.g., Posedel et al., 2011). In one study, participants were native speakers of Japanese who had lived in the United States for at least 6 months (Slevc & Miyake, 2006). They took tests of music aptitude, perception and production of English speech sounds, and knowledge of English words and syntax. Music aptitude was correlated with each of these second-language abilities. When the researchers held constant possible confounding variables (i.e., age of arrival and time spent in the United States, experience with English, and phonological short-term memory), only the association between music aptitude and facility with English sounds remained significant. Unfortunately, the authors did not control for music training, which is associated with music aptitude and with performing well.
on a variety of tests, linguistic or otherwise. Thus, associations between aptitude and second-language abilities could have stemmed from individual differences in music training rather than aptitude (Schellenberg & Peretz, 2008).

Other researchers have reported that Finnish children (Milovanov, Huotilainen, Valimäki, Esquef, & Tervaniemi, 2008; Milovanov et al., 2009) and adults (Milovanov, Pietilä, Tervaniemi, & Esquef, 2010) who perform well on a test of music aptitude also tend to have good pronunciation skills in English (for review, see Milovanov & Tervaniemi, 2011). For the adults, the association between aptitude and pronunciation remained evident when individual differences in intelligence, music training, and the ability to discriminate English phonemes were held constant. For the children, however, music aptitude was associated with the visuo-spatial portion of intelligence tests. For tests of general and verbal intelligence, the authors reported simply that associations with aptitude fell short of significance ($p > .05$; see also Milovanov et al., 2007), and they made no attempt to control for intelligence in the analyses. In short, only a few studies have provided evidence that music aptitude and language abilities rely on shared mechanisms distinct from general intelligence (Barwick et al., 1989; Lamb & Gregory, 1993; Milovanov et al., 2010).

B. Music Aptitude and Mathematics

One common belief among the public and some scholars is that music is inherently linked with mathematics (e.g., Fauvel, Flood, & Wilson, 2006). There are some sensible reasons for considering the two domains to be associated. Consider tone durations in music: 1 whole note is equivalent to 2 half notes, 4 quarter notes, 8 eighth notes, 16 sixteenth notes, and so on. Now consider pitch: an octave is created by doubling the fundamental (lowest) frequency in cycles per second. Moreover, complex tones such as those produced by singing or musical instruments have harmonics that are integer multiples of the fundamental frequency. Nevertheless, it does not follow that mathematical abilities are associated with music abilities simply because music has mathematical properties.

Music aptitude has been shown to be positively associated with basic arithmetic abilities in 4- but not 5-year-olds (Anvari et al., 2002). Music aptitude in first, second, and third graders was also shown to correlate with scholastic achievement in mathematics, but an association similar or larger in magnitude extended to achievement in reading (Hobbs, 1985). Moreover, when children were given standardized IQ tests that had many subtests with one measuring numerical abilities, correlations with aptitude were similar across subtests (Lynn, Wilson, & Gault, 1989). In one instance, the correlation between music aptitude and quantitative abilities was actually lower than the correlation between aptitude and verbal or nonverbal abilities (Phillips, 1976). Convincing evidence of a “special” link between music aptitude and mathematics requires that the association remains evident when general intelligence is held constant. We are unaware of any such findings. Moreover, mathematicians with doctoral degrees are not any more musical than similarly qualified scholars from the humanities (Haimson, Swain, & Winner, 2011).
C. Music Aptitude and General Cognitive Abilities

There is plenty of evidence that music aptitude is associated with general intelligence in childhood (Doxey & Wright, 1990; Hobbs, 1985; Lynn et al., 1989; Norton et al., 2005; Phillips, 1976; Rainbow, 1965; Sergeant & Thatcher, 1974). In one study of 10- and 11-year-olds, several tests were used to measure music aptitude and IQ (Sergeant & Thatcher, 1974). Each IQ test was correlated positively with each aptitude subtest. When 10-year-olds were given standardized IQ tests as well as tests of music aptitude, principal components analysis revealed that a one-factor solution provided a good account of the correlations among the various measures (Lynn et al., 1989). In other words, music aptitude may be a surrogate measure for general intelligence. Although other variables such as creativity, socioeconomic status (SES), and music ability and experience are associated with aptitude, intelligence continues to be associated with aptitude when these other variables are held constant (Doxey & Wright, 1990; Rainbow, 1965). Even in adulthood, basic pitch and temporal discrimination abilities are associated with intelligence (e.g., Helmbold, Troche, & Rammsayer, 2006).

Thus, associations between music aptitude and specific aspects of cognition (e.g., spatial abilities: Nelson, Barresi, & Barrett, 1992; working memory: Wallentin et al., 2010) may be a by-product of the association between aptitude and general intelligence. In line with the view that music aptitude is a marker of general intelligence, individuals with intellectual disabilities (for review see Hooper, Wigram, Carson, & Lindsay, 2008) or learning disabilities (Atterbury, 1985) tend to perform poorly on tests of music aptitude. Moreover, among individuals with mental retardation, as the degree of cognitive disability increases, music aptitude decreases (Braswell, Decuir, Hoskins, Kvet, & Oubre, 1988). Children with Williams syndrome are often considered to be an exception, with low IQs but reportedly good music abilities. Nevertheless, their music aptitude is well below norms although not as poor as their spatial abilities or overall IQ (Don, Schellenberg, & Rourke, 1999).

The association between music aptitude and cognitive abilities extends beyond IQ testing to performance in school (Good, Aggleton, Kentridge, Barker, & Neave, 1997; Gordon, 1969; Harrison, Asmus, & Serpe, 1994; Hobbs, 1985; Rainbow, 1965). In fact, the correlation between music aptitude and academic achievement can be substantially higher than the correlation between aptitude and IQ (Hobbs, 1985). Although these findings do not call into question associations between aptitude and general intelligence, they suggest that other factors that are associated with aptitude, such as SES or creativity (Doxey & Wright, 1990; Rainbow, 1965; Sergeant & Thatcher, 1974), or variables that have yet to be tested (e.g., personality), play an additional role in the association between aptitude and academic achievement.

Associations between music aptitude and general intelligence or academic achievement belie proposals of modularity for music (Peretz, 2009; Peretz & Coltheart, 2003) and the notion that music ability represents a distinct intelligence (Gardner, 1983, 1999). Rather, the evidence suggests that high-functioning children perform well on tests of music aptitude, just as they perform well on many tests of
cognitive abilities (Carroll, 1993). Some doubts about this conclusion arise, however, because correlations between aptitude and general intelligence are often small (Drake, 1954), primarily because some high-IQ individuals perform poorly on tests of aptitude (Sergeant & Thatcher, 1974). Indeed, about 4% of the population may have amusia, performing poorly on tests of aptitude but with normal IQ, hearing abilities, and exposure to music (Peretz, 2008, Chapter 13, this volume; Stewart, 2008, 2009). These individuals are usually diagnosed with the Montreal Battery of Evaluation of Amusia (Peretz, Champod, & Hyde, 2003), a test of aptitude that is designed to identify individuals with particularly low levels of aptitude. Amusia appears to be the consequence of a pitch-processing deficit (e.g., Hyde, Zatorre, & Peretz, 2011; Peretz et al., 2002), with affected individuals unable to discriminate small changes in pitch but with intact temporal processing (Hyde & Peretz, 2004). Although one study reported that amusics exhibit additional deficits on spatial tasks (Douglas & Bilkey, 2007), this finding does not appear to be reliable (Tillmann et al., 2010). Musical savants, with good musical abilities and atypical cognitive development (e.g., autism or mental retardation), also represent a notable exception to the rule that music aptitude is tightly coupled with general intelligence (Treffert, 2009).

In sum, although music aptitude is typically associated positively with general intelligence, some individuals with normal intelligence score poorly on tests of music aptitude because of a selective deficit in pitch perception, and some with low intelligence show high levels of musicality. High levels of music aptitude may also facilitate the acquisition of a second language and the ability to read.

III. Cognitive Abilities after Listening to Music

Does music listening enhance cognitive performance immediately after listening? The typical procedure used to answer this question is to have people listen to music for about 10 minutes and then complete a short test of cognitive abilities. Cognitive performance is compared with performance on another version of the same or a similar test in a control condition that substitutes listening to music with sitting in silence or listening to another auditory stimulus. Because it is relatively easy to assign participants randomly to one condition or another in a between-subjects design, or to counterbalance order of conditions in a within-subjects design, inferences of causation are considerably easier for these studies than for the question of whether music aptitude or training is associated with cognitive abilities.

The procedure just described may seem artificial, however, because people regularly do some sort of task (e.g., reading, homework, driving) while they listen to music. Indeed, interest in cognitive performance after music listening was inspired not by ecological validity but by provocative findings published in the early 1990s, which implied that visuospatial abilities could be improved after listening to music composed by Mozart (Rauscher, Shaw, & Ky, 1993). Since then, there have been numerous attempts to replicate the so-called “Mozart effect” and to determine
whether the effect, if it exists, is limited to music composed by Mozart, to music in general, or to tests of visuospatial abilities (for a review see Schellenberg, 2012).

### A. The Mozart Effect

In the original study, undergraduates completed three tests of visuospatial reasoning, one test after each of three different listening experiences (Rauscher et al., 1993). The tests were three different subtests from a test of intelligence, each standardized to have the same mean and standard deviation. The experimental condition involved listening to 10 minutes of a Mozart sonata (K. 448), whereas the control conditions involved 10 minutes of either listening to relaxation instructions or sitting in silence. The design precluded testing for an interaction between listening condition and test because each participant had data for only three of the nine possible combinations of condition and test. Moreover, in order to compare the three listening conditions (i.e., the authors’ goal), the three tests had to be assumed to be measuring the same thing so that each participant had a score in each of the three conditions. Fortunately, performance on the visuospatial tasks was correlated; participants who performed well on one test also tended to perform well on the other two. Because performance across tests was significantly higher in the Mozart condition than in the other two conditions, this finding became known as the “Mozart effect.”

In the first published attempt to replicate the effect, researchers administered a test of general intelligence after participants listened to 10 minutes of the Mozart piece used by Rauscher et al. (1993) or 10 minutes of dance music, or after they sat in silence for 10 minutes (Stough, Kerkin, Bates, & Mangan, 1994). In contrast to the original findings, performance did not vary across conditions. In fact, differences between conditions were so small that the null result did not appear to stem from a lack of statistical power.

Perhaps even more frustrating to those who wanted to believe in the effect were subsequent studies published by the original research team, who backtracked on the original claim of a link between listening to Mozart and spatial reasoning (or abstract reasoning, spatial IQ, or IQ scores, terms they used interchangeably in the original report). In order to explain subsequent failures to replicate, the effect was now said to be limited to tests of “spatial-temporal” reasoning, based on predictions from a theory (i.e., the Trion model; Leng & Shaw, 1991) that posits a neurological priming effect between listening to Mozart and spatial-temporal abilities (Rauscher & Shaw, 1998; Rauscher, Shaw, & Ky, 1995). The Trion model has yet to be recognized in neuropsychology or neuroscience textbooks or in any scholarly articles outside of those testing the Mozart effect, and spatial-temporal reasoning is not considered to be an ability distinct from other visuospatial abilities in established models of intelligence (Carroll, 1993; Gardner, 1983, 1999; Linn & Petersen, 1985). Object Assembly, the single subtest in the Wechsler Intelligence Scale for Children (WISC) and the Wechsler Adult Intelligence Scale (WAIS) that is considered to measure spatial-temporal ability (e.g., Rauscher et al., 1997), correlates strongly with other visuospatial subtests from the same battery (e.g., Block Design, Picture Completion), which is what makes it possible for them to be
combined to form the Perceptual Reasoning index, a measure of general visuospatial ability.

**B. Meta-analyses**

The first meta-analysis of the Mozart effect examined 16 published studies with more than 714 participants in total (Chabris, 1999). Each study compared performance between participants who took a cognitive test after listening to Mozart and other participants who took the same test after sitting in silence or listening to relaxation instructions. The tests included measures of abstract reasoning and a test that Rauscher considers to measure spatial-temporal ability—the Paper Folding and Cutting (PF&C) test. The PF&C test requires participants to envision how a piece of paper that has been folded and cut will look after it is unfolded. Results of the meta-analysis showed no difference in performance after listening to Mozart versus sitting in silence regardless of the test administered. The results did, however, show significantly higher performance after listening to Mozart versus listening to relaxation instructions, particularly for the PF&C test. Because relaxation instructions aim to calm listeners and reduce their arousal levels, Chabris speculated that better performance in the Mozart condition could be attributed to participants being more aroused after listening to Mozart.

A meta-analysis published the next year reached a different conclusion (Hetland, 2000b). Hetland examined 36 independent experiments of visuospatial abilities with more than 2400 participants in total. In contrast to Chabris (1999), Hetland included published and unpublished data, experiments with within- as well as between-subject designs, and experiments with different control conditions (i.e., silence, relaxation instructions, other kinds of music) that were contrasted with listening to classical music (i.e., composed by Mozart and others). The meta-analysis revealed that the observed advantage for the classical music conditions was driven by studies that used spatial-temporal outcome measures. Hetland’s other conclusions were that the effect was (1) evident for classical music whether or not it was composed by Mozart, (2) larger when relaxation instructions rather than silence were used as a control condition, (3) independent of the gender of the listener, and (4) stronger in Rauscher’s laboratory than in other laboratories.

A more recent meta-analysis examined 39 studies (> 3000 participants in total) that contrasted listening to the original Mozart piece (K. 448) with listening to a nonmusical stimulus or sitting in silence (Pietschnig, Voracek, & Formann, 2010). An overall Mozart effect was evident. Although the authors described the mean effect size \( d = 0.37 \) as “small,” it was closer to levels typically considered to be “moderate” \( d = 0.5 \) rather than “small” \( d = 0.2 \) in magnitude (Cohen, 1988). As in Hetland’s (2000b) meta-analysis, the effect size was similar when another piece of music (composed by Mozart or others) was contrasted with silence or a nonmusical stimulus, and it tended to be larger when tested by Rauscher (or Rideout) compared with other researchers. In contrast to Hetland, the effect size was similar whether researchers used a spatial-temporal test such as the PF&C test or other tests of visuospatial abilities.
In sum, the most recent and comprehensive meta-analysis confirms that there is a Mozart effect, although the effect is also evident for other pieces of music and independent of whether the cognitive test measures spatial-temporal abilities. Why the effect is larger in studies conducted by two particular researchers compared with others remains unknown. Regardless, because three separate meta-analyses found some evidence of the effect (Chabris, 1999; Hetland, 2000b; Pietschnig et al., 2010), the question of whether the effect exists is no longer in contention. However, as we have argued elsewhere (e.g., Schellenberg, 2005, 2006a, 2012), the effect seems particularly elusive when participants are tested in groups rather than individually (e.g., Črnčec, Wilson, & Prior, 2006; Steele, Bass, & Crook, 1999; Steele, Dalla Bella, et al., 1999), presumably because listening intently to recorded music or sitting in silence in a group context is an unusual activity.

C. Arousal and Mood

Schellenberg and his colleagues (Husain, Thompson, & Schellenberg, 2002; Nantais & Schellenberg, 1999; Schellenberg & Hallam, 2005; Schellenberg, Nakata, Hunter, & Tamoto, 2007; Thompson, Schellenberg, & Husain, 2001) sought to determine whether the Mozart effect, when evident, was a consequence of arousal and mood. The basic idea was that any stimulus that improves how a person feels can in turn improve how they perform on a cognitive task. The advantage of this hypothesis is that it explains the Mozart effect with two links that are well established in the psychological literature.

Consider links between music listening and emotional responding. People choose to listen to music because of the way it makes them feel (e.g., Juslin & Västfjäll, 2008; Lonsdale & North, 2011; Sloboda, 1992). Calming music changes cortisol levels (Flaten, Asli, & Simonsen, 2006) and blood pressure (Triller, Erzen, Dub, Petrinic-Primožič & Kosnik, 2006), and listening to music reduces anxiety in medical contexts (Bare & Dundes, 2004; Cooke, Chaboyer, Schluter, & Hiratos, 2005; Pelletier, 2004; Weeks & Nilsson, 2010). Music listening also facilitates falling asleep (Field, 1999), and it increases levels of sedation for patients in intensive care (Dijkstra, Gamel, van der Bijl, Bots, & Kesecioglu, 2010). Listening to one’s preferred music can even reduce perceived pain levels after surgery (Ebneshahidi & Mohseni, 2008). Music’s ability to calm the listener is one reason why music therapy often yields positive effects (Gold, Voracek & Wigram, 2004). Thus, it is clear that music listening can change one’s emotional state.

There is also an abundance of evidence that feelings influence cognitive performance (e.g., Cassady, Mohammed, & Mathieu, 2004; Isen & Labroo, 2003, O’Hanlon, 1981). Isen (2009) illustrated how positive affect enhances cognitive abilities, including decision making, problem solving, social interaction, and thought processes in general. Positive affect is associated with increases in dopamine levels, which may improve cognitive flexibility (Ashby, Isen, & Turken, 1999). Effects of emotional state on cognition are evident even with small increases in positive affect, such as those that occur when receiving a small gift or watching a comic film. For example, in a problem-solving task that required participants to
solve Duncker’s (1945) candle problem, participants who saw a brief comic film just before performed better than control participants (Isen, Daubman, & Nowicki, 1987). Similarly, participants who were given a small bag of candy performed better than controls on a test of remote associates (Isen et al., 1987). Moreover, negative affect (e.g., boredom) impairs cognitive performance (Cassady et al., 2004; O’Hanlon, 1981). Thus, links between arousal and/or mood and cognitive performance are also well established.

In the first of a series of studies, Nantais and Schellenberg (1999) had participants complete one of two sets of PF&C items on two different visits to the laboratory, once after listening to Mozart (K. 448) for 10 minutes and once after sitting in silence. The order of the conditions (Mozart then silence or silence then Mozart) and the two sets of PF&C items was counterbalanced, with an equal number of participants in each of the four cells. PF&C performance was better after listening to Mozart than after sitting in silence, thus replicating the Mozart effect. This result was not surprising because the control condition involved sitting in silence while staring at a computer monitor for 10 minutes, which was unlikely to put participants in an optimal state of mind for any sort of task. A separate group of participants was tested identically but the Mozart music was replaced with a piece composed by Schubert from the same CD, performed by the same pianists with the same production values. For these participants, a Schubert effect (i.e., better PF&C performance after listening to Schubert than sitting in silence) was evident and similar in magnitude to the Mozart effect. Again, adverse aspects of the control condition virtually guaranteed the outcome.

For a third group of participants, Nantais and Schellenberg (1999) used the same experimental design but they compared PF&C performance after listening to Mozart or a narrated story written by Stephen King. The story was chosen because it was an auditory stimulus that changed over time (like the Mozart piece) and it was likely to be as enjoyable as listening to Mozart for a sample of college freshmen. Accordingly, the authors predicted that both Mozart and King should enhance cognitive performance equally, and this prediction was confirmed. After completing the experiment, participants were asked which listening experience they preferred. Approximately half preferred the music; the other half preferred the story. When the data were reanalyzed as a function of preference and condition, performance was better in the preferred than in the nonpreferred conditions. In other words, there was a Mozart effect for participants who preferred the music, but a Stephen King effect for participants who preferred the story.

Thompson et al. (2001) used the same basic design. Each participant completed a version of the PF&C task after listening to music and sitting in silence. The music was the same Mozart piece used earlier (K. 448) for half of the participants, and Albinoni’s Adagio for the other half. The Adagio is a classic example of sad-sounding music, written in a minor key with a slow tempo. Measures of arousal and mood were administered before and after the listening experiences. The hypotheses were that (1) there would be no Albinoni effect because the piece sounds so somber, (2) a Mozart effect would once again be evident (the first movement is happy-sounding with a fast tempo and in a major key), and (3) the
advantage for Mozart over silence would be eliminated after controlling for changes in arousal or mood. The results were consistent with these predictions. Because the Albinoni piece is relatively well known and considered to be a quintessential piece of sad-sounding music, however, it could have evoked particular associations with sad events, such that these associations rather than the music were the source of the null effect.

Accordingly, Husain et al. (2002) conducted an additional experiment in which they manipulated the tempo and mode of the same piece (Mozart K. 448) using MIDI. As noted, happy-sounding music tends to be fast-tempo and in a major key, whereas sad-sounding music tends to be slow and minor (Hunter & Schellenberg, 2010). Each participant completed the PF&C task after listening to one version of the Mozart piece: fast/major, slow/major, fast/minor, or slow/minor. Arousal and mood were measured before and after listening. PF&C performance was better after the fast and major-key versions. Arousal levels were improved after hearing the fast versions, whereas moods were improved after hearing the major-key versions. As predicted, changes in arousal and mood accounted for the bulk of the variance in PF&C scores.

If effects of music on cognition are determined by music’s emotional impact, then the type of music that is the most effective in this regard should depend critically on the particular sample of listeners. In the next study (Schellenberg & Hallam, 2005), more than 8,000 10- and 11-year-olds living in the United Kingdom were recruited from approximately 200 schools to participate in a study of the Mozart effect, which was coordinated by the British Broadcasting Corporation (BBC). At each school, students were assigned to one of three rooms at exactly the same time. In one room, they heard pop music on BBC Radio 1, including a song by the band Blur. In a second room, they heard a Mozart piece over BBC Radio 3. In the third room, they listened to Susan Hallam discuss the experiment on BBC Radio 5. After listening to the radio, they completed two tests of spatial abilities. According to the arousal and mood hypothesis, the pop songs should be the most likely to put the children in an optimal emotional state for the cognitive tests. On the simpler of the two spatial tests, no differences among groups emerged. On the more difficult test, however, performance was indeed best for the children who heard pop songs. In other words, a Blur effect was evident for 10- and 11-year-olds living in the United Kingdom.

Taking this approach even further, Schellenberg et al. (2007) tested the creative abilities of 5-year-old Japanese children. Recall that the arousal and mood hypothesis does not give special status to music listening. Any experience that changes the participant’s emotional state can influence cognitive performance. Moreover, the hypothesis extends beyond spatial abilities to cognitive performance construed broadly, including creativity, which is considered to be one aspect of cognition (e.g., Sternberg, 2009). In an initial (baseline) session, children were given a piece of paper and 18 crayons and asked to draw whatever they liked. They subsequently made another drawing after a musical experience: listening to Mozart (K. 448), Albinoni’s Adagio, or children’s playsongs, or singing children’s playsongs. The hypothesis was that creativity would be enhanced after singing or listening to the
playsongs. The dependent measures were drawing times and ratings made by adults about each child’s pair of drawings in terms of creativity, energy, and technical proficiency. In line with predictions, drawing times increased more from baseline in both playsong groups than in the Mozart and Albinoni groups. Ratings of creativity, energy, and technical proficiency also increased relative to baseline for children who heard or sang playsongs, but these ratings decreased for children who heard Mozart or Albinoni. In short, a playsong effect was evident for Japanese 5-year-olds.

In another experiment that examined whether the Mozart effect extends beyond measures of spatial ability, college freshmen were tested on measures of processing speed or working memory after they heard Mozart or Albinoni (Schellenberg et al., 2007). Each student came to the lab twice, both times completing one test after one listening experience. The order of the tests and the musical pieces was counterbalanced. Arousal and mood were measured before and after the music listening. At the first testing session, music listening led to inconsistent changes in arousal and mood and there were no differences between conditions on either cognitive test. At the second session, the data were consistent with predictions. Arousal increased after listening to Mozart but decreased after listening to Albinoni; similarly, mood improved after listening to Mozart but declined after listening to Albinoni. Although performance on both cognitive tests was better after listening to Mozart, the comparison was significant only for the test of processing speed. When these data are considered jointly with those from the British children (Schellenberg & Hallam, 2005), it appears that effects of emotional state on cognition may indeed be greater on some tasks (e.g., more difficult tests, tests of visuospatial abilities or processing speed) than on others (e.g., easier tests, tests of working memory; see also Rauscher et al., 1995; Steele, Ball, & Runk, 1997).

In any event, there is no compelling evidence of a special link between listening to Mozart (or to any Classical music) and visuospatial (or spatial-temporal) abilities. Rather, listening to music is an effective way to improve one’s emotional state, and how one feels can influence cognitive performance quite generally. As an aside, very few published studies have examined mathematical performance after listening to music. Although there was improvement in one study from pretest to posttest on a measure of mathematical ability, such improvement was similar whether participants listened to Mozart, Bach, or ocean sounds in the interim (Bridgett & Cuevas, 2000). In another study (Jaušovec & Habe, 2003), performance on a mathematical task was slightly (but not significantly) worse after listening to Mozart than after sitting in silence.

IV. Background Music and Cognitive Abilities

Is cognitive performance affected while music plays in the background? Consider listening to music while driving a car. Although there is evidence that background music negatively affects driving (re: wheel movements, traffic violations, perceived and actual speed; Brodsky, 2002; Konz & McDougal, 1968),
many drivers listen to music and consider it to be less distracting than conversation (Dibben & Williamson, 2007). Background music also influences athletic performance (Pates, Karageorghis, Fryer, & Maynard, 2003), altruistic behavior (North, Tarrant, & Hargreaves, 2004), drinking rate (McElrea & Standing, 1992), self-disclosure (Jensen, 2001), and sleep quality (Tan, 2004), yet listeners are often unaware of its effects.

Because of the widespread use of background music in commercial environments, much of the available research has examined consumer behavior. The typical finding is that consumers unwittingly alter their behavior to fit with the music. For example, French wine outsells German wine when stereotypical French music is played in a supermarket, but German wine outsells French wine when stereotypical German music is played (North, Hargreaves, & McKendrick, 1997, 1999). Similarly, diners tend to order food that matches the ethnicity of background music that is played in a restaurant (Yeoh & North, 2010). Diners (North, Shilcock, & Hargreaves, 2003) and wine buyers (Areni & Kim, 1993) tend to spend more in a shop when classical rather than pop music is played in the background, presumably because classical music is associated with wealth and affluence. Because it is relatively simple to manipulate exposure to background music, it is clear that background music causes changes in behavior.

The question of whether background music affects performance on tests of cognitive abilities is particularly pertinent for students who study while listening to music (Patton, Stinard, & Routh, 1983) or with media of some sort playing in the background (Beentjes, Koolstra & van der Voort, 1996). Even if students turn the music on or off based on their perception of its degree of distraction (Kotsopoulou & Hallam, 2010), they cannot always predict its effects (Alley & Greene, 2008).

A. Emotional Responses and Cognitive Capacity

Early psychological research documented some instances in which background music interfered with cognitive tasks like reading comprehension (Fendrick, 1937), but only when the music had vocals (Henderson, Crews, & Barlow, 1945). In another instance, background music had no effect on reading whether or not it had vocals (Freeburne & Fleischer, 1952). In yet another instance, reading comprehension among eighth and ninth graders improved in the presence of background music (Hall, 1952). These inconsistent findings remain emblematic of contemporary studies, as evidenced by a recent meta-analysis reporting an overall null effect of background music on cognitive abilities (Kämpfe, Sedlmeier, & Renkewitz, 2011). More detailed analyses allowed Kämpfe et al. to conclude that background music has small detrimental effects on reading and memory.

The conflicting findings likely stem directly from the nature of background music. On the one hand, because the term “background music” implies that listeners are doing two things at the same time, cognitive limitations are likely to play a role, which could lead to decrements in performance on the primary task. On the other hand, music often improves listeners’ emotional states, which can lead to better performance on tests of cognitive abilities. In the context of the
Mozart effect, the arousal and mood hypothesis describes mechanisms through which music improves performance on cognitive tasks after enjoyable music has stopped playing (Husain et al., 2002; Thompson et al., 2001). These mechanisms should apply similarly to background music although the direction of the effect need not always be positive. For example, sad- or aggressive-sounding music that causes less than optimal arousal levels or moods might affect cognitive performance negatively.

According to the cognitive capacity model (Kahneman, 1973), different cognitive processes draw from the same limited pool of resources. When multiple tasks are performed simultaneously they can overtax available resources, leading to cognitive interference (Norman & Bobrow, 1975). For example, the sound of a television in the background negatively influences performance on a range of cognitive tasks because it overloads cognitive capacity, particularly when participants are instructed to attend to the soundtrack while completing the task (Armstrong & Greenberg, 1990). More generally, when music is presented during a cognitive task, it may compete for resources and impair efficiency. For example, surgeons learning a new procedure perform worse in the presence of background music (Miskovic et al., 2008). In a virtual driving task, participants perform worse and rate concurrent tasks as more difficult in the presence of highly arousing background music (North & Hargreaves, 1999).

Whether cognitive load is overtaxed also depends on depth of processing (Lavie, 2005; Lavie, Hirst, de Fockert & Viding, 2004) and type of information. Models of working memory posit that the type of input affects cognitive load (Baddeley, 2003). According to Baddeley’s (1986) model, working memory comprises a phonological loop, a visuospatial sketchpad, and an executive control system. Because background music is an auditory stimulus, it uses available resources from the phonological loop. Accordingly, when the primary task also uses the phonological loop, such as when a participant is rehearsing verbal information during reading, the likelihood of interference should increase. In general, if concurrent stimuli are processed through the same channels, they may overload cognitive capacity, whereas stimuli processed in separate channels are less likely to do so. In line with this view, background music may not affect reaction times when it accompanies visual materials, but it can slow people down when it accompanies an audiovisual presentation of the same information (Brünken, Plass & Leutner, 2004).

The irrelevant sound effect refers to instances when a concurrent auditory stimulus interferes with other working-memory processes. Because sounds like background music are processed obligatorily (i.e., we can close our eyes but not our ears), they can disrupt other working-memory processes that track changes over time (Banbury, Macken, Tremblay & Jones, 2001; Jones, Macken & Nicholls, 2004). Background sounds that change over time impair serial recall, whether the sound is played at a high or low volume, and whether it is instrumental or vocal, but not if the irrelevant sound is constant (unchanging) pink noise (Ellermeier & Hellbrück; 1998). In fact, performance on the primary task improves as the ratio of pink noise to irrelevant sound (i.e., masking) increases.
B. Background Music and Mathematics

In some contexts with some populations, background music improves performance on tests of mathematics or arithmetic. In one study, emotionally disturbed 9- and 10-year-olds with normal IQs completed an arithmetic task in silence and then with calming music played in the background (Hallam & Price, 1998). This procedure was repeated a week later with the music and silence conditions in reverse order. For each child, the researchers measured performance on the test and the number of disruptive incidents each student initiated. Arithmetic performance was significantly better in the music condition than in the silence condition. Because math performance was negatively correlated with the number of incidents of disruptive behavior, the music appears to have had a relaxing influence, which led to better behavior, increased focus, and consequently better arithmetic performance.

In another study of hyperactive children 7 to 11 years of age (Scott, 1970), arithmetic performance was again better in the presence of background music than in silence. The music consisted of recordings from The Beatles, which shows that positive effects of background music on math performance are not limited to calming music. Rather, the most effective music likely depends on the particular population, as it does when cognitive performance is tested after music listening (Schellenberg & Hallam, 2005; Schellenberg et al., 2007).

In another study, typically developing 10- and 11-year-olds completed an arithmetic task with calming music played in the background (Hallam, Price, & Katsarou, 2002). Compared with a group who did the task in silence, the music group completed significantly more problems but the number of problems solved correctly did not vary between groups. Thus, calming music enhanced the speed but not the quality of work. Moreover, when Bloor (2009) administered a mathematics test to 10-year-olds, performance was poorer in the background music than in the silence condition. Because the music was a piece composed by Mozart, the children may have found it distracting or annoying rather than arousing or calming. Again, the choice of background music and the particular population are bound to influence whether improvements or decrements in cognitive performance are observed. In one study (Wolfe, 1983), familiar instrumental music had no effect on college students’ mathematics performance regardless of volume, even when the music was very loud and considered by the students to be distracting.

C. Background Music and Memory

Evidence of the influence of background music on memory is similarly confusing. In one study, Salamé and Baddeley (1989) varied the type of sound presented during a digit-recall task. Their participants were instructed to memorize a series of nine digits presented individually on a computer screen, after which they reproduced the digits in writing. Both vocal and instrumental background music impaired performance when compared with silence, with vocal music having a stronger negative impact. Instrumental music impaired performance more than “modulated” noise (i.e., with amplitude varied in a speechlike way). Other studies
have found that visual serial recall is impaired when a sequence of varying tones or syllables is played in the background but not when the tones or syllables are simply repeated (Jones & Macken, 1993). As noted, the amount of acoustic change in the background music matters, with greater change using more of the capacity of the phonological loop. Thus, continuous noise masks irrelevant office sounds and improves performance on verbal serial recall, but music does not (Schlittmeier & Hellbrück, 2009). Because verbal-recall tasks require rehearsal and maintenance of verbal information, music often impairs performance.

Nevertheless, when information is presented in the context of an audiovisual immersive environment, memory for the pertinent facts can be improved when the narration is presented in the presence of background music compared with silence (Richards, Fassbender, Bilgin, & Thompson, 2008). Conflicting findings come from a study that measured knowledge acquired from a multimedia presentation. Background music led to poorer performance compared with silence, and music with additional sound effects led to the poorest retention of all (Moreno & Mayer, 2000). These results highlight the struggle of educators who try to use multimedia to engage learners without distracting them from the content or overloading the senses. Placement of the music and how it fits with the primary task also matters. For example, when recall of the content of a movie scene is tested, performance is better when music of an incongruent mood precedes the content; music of a congruent mood results in better performance when it is presented during the content (Boltz, Schulkind, & Kantra, 1991). In both cases, music appears to facilitate memory when it draws attention to the desired content (through a contrast effect or by complementing the content, respectively) without competing with it.

Other factors that influence memory include individual differences and the music itself, which may explain why background music facilitates memory in some contexts but not in others. For example, de Groot (2006) tested memory for nonwords using a paired-associates task. A native-language (Dutch) word was paired with a nonword during learning. During the subsequent test phase, participants were provided with the nonword and asked to provide the corresponding Dutch word. Background music consisted of an excerpt from a Brandenburg Concerto, which was presented during the learning phase for half of the participants. There was improvement in recall of Dutch words for those who learned with the background music compared with those who learned in silence, but the effect was evident only when the items (not the participants) were considered to be the experimental unit. In other words, because some but not all of the participants exhibited the effect, a positive result was observed only when performance of the entire group was examined as a function of the test items.

The effect of background music on an associative-learning memory task can also depend on personal study habits. Students who typically study with background music may perform no worse on the task when background music is presented, whereas students who usually study in silence can be affected negatively by the music, particularly when it has lyrics (Crawford & Strapp, 1994). Other evidence suggests that high-arousal music is more disruptive to memory than relaxing music, both for children (Hallam et al., 2002) and for adults (Cassidy & MacDonald, 2007).
The variety of outcomes is consistent with the notion that music engages multiple processes, some of which facilitate memory and others that compete with it.

Some memory experiments vary background music from learning to test, treating the musical stimulus as the “context” in which learning occurred. From this perspective, performance should decline with a shifting context because the new context does not prime the learned content. In one study, verbal memory for a word list was worse when the music changed from learning to test compared with when it stayed the same (Standing, Bobbitt, Boisvert, Dayholos, & Gagnon, 2008). Balch, Bowman, and Mohler (1992) found that changing the musical context impaired immediate recall but it had no effect on recall after a 48-hour delay. In one instance, a change in tempo between learning and test negatively affected recall, but changing other features of the music or removing the music at test had no effect (Balch & Lewis, 1996). In another instance, background music that varied in tempo and consonance was presented while participants saw individual nonsense words presented one at a time on a computer monitor (Jäncke & Sandmann, 2010). Their task was to identify whether each word had appeared previously. Although the test was clearly a measure of verbal learning and memory, the presence of background music did not affect performance. In sum, whether background music enhances or hinders memory depends on multiple factors.

D. Background Music and Reading Comprehension

Effects of background music on reading comprehension are particularly important to understand because students often listen to background music while they study (Patton et al., 1983), and studying almost always involves reading. Even without background music, understanding what we read is cognitively demanding because it requires synchronization of automatic processes like letter identification and semantic access, as well as attention-demanding processes like inference generation and text elaboration (Walczyk, 2000). Models of reading comprehension assume a role for cognitive-capacity limitations (Carretti, Borella, Cornoldi, & De Beni, 2009; Kintsch & van Dijk, 1978). The obligatory processing of music may interfere directly with reading and overload working memory (both types of information are processed in the phonological loop), especially if the reading task is difficult.

From this perspective, it is easy to understand why attending to background music interferes with performance on reading-comprehension tasks (Madsen, 1987). Because background music is by definition a secondary stimulus, however, we may not attend to it much if at all. In one study of college students, researchers varied the intensity and tempo of background classical music during reading, after which participants answered multiple-choice questions about the passage’s content (Thompson, Schellenberg, & Letnic, 2011). When the background music was fast and loud, and thus difficult to ignore, reading comprehension declined relative to baseline (no music). Because the other conditions (slow/soft, slow/loud, fast/soft) did not influence performance compared with baseline, it appears that participants were able to focus on the reading task when the music was below a perceptual threshold. Another study of seventh to eighth graders presented familiar pop songs as background music
during the reading and test phases (Anderson & Fuller, 2010). Comprehension declined compared with the control (silent) condition, and the effect was evident in 75% of participants. Relatively benign effects in one instance (Thompson et al., 2011) compared with a strong effect in another instance (Anderson & Fuller, 2010) could be due to differences in age, musical styles, the presence of vocals, different tests of reading comprehension, how much participants liked the music, and/or presentation of background music during the test phase.

In situations such as a noisy cafeteria, music can actually mask distractions when reading. In one study, university students read faster and remembered more content when the background music in a cafeteria consisted of fast-tempo classical music compared with conditions in which the music was slower or there was no music at all (Kallinen, 2002). This finding conflicts with the results of a study of working memory (Schlittmeier & Hellbrück, 2009), which found that music did not mask office noises. The task in the office context involved serial recall, however, which differs markedly from sustained periods of reading during which participants may reread certain sections. Regardless, whether music distracts the reader or prevents distraction is undoubtedly influenced by the context. In an otherwise silent room, background music might impair reading comprehension; in public environments, background music might mask distracting and unpredictable noise. These conflicting factors help to explain why several studies have failed to find effects of background music on reading comprehension among samples of children (Bloor, 2009; Furnham & Stephenson, 2007), high-school students (Pool, Koolstra, & van der Voort, 2003), or adults (Boyle & Coltheart, 1996; Freeburne & Fleischer, 1952).

E. Background Music and Individual Differences

The results of de Groot (2006) suggest that individual differences influence whether background music facilitates or impairs cognitive processing. One dimension that may be particularly important is personality. For example, introverts and extroverts perform differently on tests of memory and reading in the presence of pop music or silence (Furnham & Bradley, 1997). On a test of immediate memory recall, background music hinders performance of both groups. For delayed-recall and reading-comprehension tasks, however, introverts who hear pop music perform worse than either introverts who are tested in silence or extroverts who are tested with music.

In some instances, both music and noise impair reading comprehension for both groups but introverts suffer more than extroverts (Furnham & Strbac, 2002). In other instances (Furnham & Allass, 1999), introverts exhibit decrements in performance on a variety of cognitive tests in the presence of simple background (pop) music compared with silence, and with complex compared with simple pop music. On tests of delayed and immediate recall, extroverts can show the exact opposite pattern (best performance with complex music, worst performance in silence). Similarly, on a Stroop task, the negative effect of high-arousing background music is exaggerated for introverts (Cassidy & MacDonald, 2007). Even on tests of general cognitive ability, although introverts and extroverts perform
similarly in silence, introverts’ performance is impaired to a greater degree in the presence of background music or noise, with the most introverted participants exhibiting the most negative effects (Dobbs, Furnham, & McClelland, 2011).

How can this difference between introverts and extroverts be explained? Personality may be associated with different levels of arousal, with introverts having higher baseline arousal levels than extroverts (Eysenck, 1967). Thus, background music leads to more optimal levels of arousal among extroverts but to overarousal among introverts. Introverts’ sensitivity to disruptive effects of background music also interacts with the difficulty of the task. On a free-recall task, both extroverts and introverts perform poorly and similarly in the presence of background music or noise compared with silence (Furnham & Strbac, 2002). On verbal tasks that involve completing a sentence or identifying antonyms or ungrammatical sentences, background music can have no effect for either introverts or extroverts (Dobbs et al., 2011). Other research on effects of personality and background music on cognitive performance has led to similarly null findings (Furnham, Trew & Sneade, 1999; Ravaja & Kallinen, 2004).

Music training is another individual-difference variable that interacts with background music in its effect on cognitive abilities. In a study that required musicians and nonmusicians to make grammaticality judgments of individual sentences, the musicians but not the nonmusicians were affected detrimentally by the presence of background piano music, particularly when the music was distorted by the inclusion of harmonically incorrect notes (Patston & Tippett, 2011). In this instance, musicians may have attended more than nonmusicians to the music, thereby making it more distracting. Nevertheless, neither group was affected by background music when completing a visuospatial task.

F. Background Music: Conclusions

The effects of background music on cognition are dependent on many factors (Hallam & MacDonald, 2009). Even for the same task, background music can facilitate, impair, or have no effect on performance. Variables that undoubtedly play a role include individual differences (e.g., personality, music training, music preferences, study habits), the type of cognitive task, the context, and the choice of background music in terms of its mood or pleasantness (Cassidy & MacDonald, 2007; Gowensmith & Bloom, 1997; Hallam et al., 2002; Stratton & Zalanowski, 1991; Ziv & Goshen, 2006), tempo (Day, Lin, Huang, & Chuang, 2009; Kallinen, 2002; Thompson et al., 2011; Wakshlag, Reitz, & Zillmann, 1982), intensity or dynamics (Beh & Hirst; 1999; Gowensmith & Bloom, 1997; Schlittmeier, Hellbrück & Klatte, 2008; Thompson et al., 2011; Turner, Fernandez, & Nelson, 1996), and whether it has vocals (Boyle & Coltheart, 1996; Martin, Wogalter, & Forlano, 1988; Salamé & Baddeley, 1989). To date, the findings are marked by inconsistency. One final illustrative example comes from a study that measured reading comprehension and memory for lists of words in the presence of singing, singing and instrumental music, instrumental music, speech, or silence (Boyle & Coltheart, 1996). Although one might expect the conditions with singing or speech to be the most disruptive on both
verbal tasks, this effect was evident for the memory task but not for the reading-comprehension task.

In sum, there are few reliable effects of background music on cognitive abilities, as Kämpfe et al. (2011) concluded from their meta-analysis. Considering emotional responses combined with cognitive load may serve as a starting point for future research. Many variables that moderate the effect of background music (i.e., the individual, the context, the task, and the music) differ from study to study, however, such that background music might affect the same class of cognitive tests in different ways. Indeed, there is no single type of “background music,” so perhaps we should not expect straightforward results. As noted by Hallam and MacDonald (2009), a complete model of effects of background music on cognitive abilities needs to consider many variables and interactions among variables, which will make it difficult to test. The best approach might be two-pronged: (1) to document general patterns that are evident in large samples of participants and (2) to identify idiosyncratic factors that qualify these general trends.

V. Music Training and Cognitive Abilities

A. Music Training and Listening

As one would expect, musically trained participants outperform their untrained counterparts on a variety of tests of music cognition. For example, young children who take music lessons for 8 months are more likely than untrained children to notice when a familiar harmonized melody ends with an anomalous chord (Corrigall & Trainor, 2009). Children (Lamont, 1998) and adults (Halpern, Kwak, Bartlett, & Dowling, 1996) with music training also show more systematic response patterns when they rate the suitability of individual test tones that follow a key-defining musical stimulus, deeming the tonic and nonscale tones, respectively, to be most and least suitable. Moreover, musically trained children (Morrongiello & Roes, 1990) and adults (Walker, 1987) are more likely than untrained individuals to map musical dimensions (i.e., pitch changes or duration) systematically onto visual dimensions (i.e., up/down or length).

Older and younger adults with music training are also better than their untrained counterparts at recognizing melodies presented in transposition (Halpern, Bartlett, & Dowling, 1995) or at an unusually fast or slow tempo (Andrews, Dowling, Bartlett, & Halpern, 1998). Because musically trained adults have better relative pitch than untrained adults, they detect smaller mistunings to individual notes of a familiar melody (Schellenberg & Moreno, 2010) or to one note of a major chord (Koelsch, Schröger, & Tervaniemi, 1999). Finally, musically trained individuals are more accurate at determining how many notes are played simultaneously in a chord (Burton, Morton, & Abbess, 1989). Although nonmusicians may perform similarly to musicians on tasks that measure implicit knowledge of Western music (Bigand & Poulin-Charronnat, 2006), musically trained individuals outperform untrained listeners on many explicit tests of music cognition.
Advantages for participants with music training extend to lower-level auditory tasks. For example, compared with untrained participants, they have finer frequency-discrimination abilities (Jakobson, Cuddy, & Kilgour, 2003; Micheyl, Delhommeeau, Perrot, & Oxenham, 2006; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Schellenberg & Moreno, 2010; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005) and they are better at detecting differences in pitch between two tones presented for very brief durations (Marie, Kujala, & Besson, 2012; Schellenberg & Moreno, 2010). They also perform better on tasks measuring temporal discrimination (Jeon & Fricke, 1997; Marie et al., 2012; Rammsayer & Altenmüller, 2006) and timbre discrimination (Chartrand & Belin, 2006). Other psychophysical research demonstrates that musically trained individuals are more accurate than untrained individuals at detecting (1) individual harmonics in a complex tone (Fine & Moore, 1993), (2) a repeated target tone in the presence of distractor (masking) tones (Oxenham, Fligor, Mason, & Kidd, 2003), and (3) a tone presented before a masking noise (Strait et al., 2010).

Performance advantages for musically trained children and adults are also evident on a variety of low-level tests of speech perception (Patel, 2011; Patel & Iversen, 2007), such as when stimuli in frequency-discrimination and temporal-discrimination tasks are speech sounds rather than tones (Bidelman & Krishnan, 2010; Chobert, Marie, François, Schön, & Besson, 2011). Trained participants are also better than untrained participants at detecting subtle changes in pitch to the final word of a sentence, as they are at detecting pitch changes to the final note of a melody (Besson, Schön, Moreno, Santos, & Magne, 2007). Indeed, 6 months of music training are sufficient to improve detection of incongruous changes in pitch that occur at the end of sentences or melodies (Moreno et al., 2009). The ability to map more complex pitch patterns in speech onto the emotions they signal is enhanced among musically trained adults and children in some instances but not in others (Lima & Castro, 2011; Thompson, Schellenberg, & Husain, 2003, 2004; Trimmer & Cuddy, 2008).

Musically trained participants are better than untrained participants at perceiving speech in noise (Parbery-Clark et al., 2009), a task that is particularly relevant to real-life listening contexts. Musically trained individuals also perform more accurately but more slowly on tests of voice discrimination (Chartrand & Belin, 2006). Presumably, they approach the task in a qualitatively different, more analytical and detailed manner. When foreign-language stimuli are used, musicians are better than nonmusicians at discriminating vowels (Sadakata & Sekiyama, 2011) and sequences of words (Marie, Delogu, Lampis, Belardinelli, & Besson, 2011). They also show advantages at learning to use pitch patterns to identify words (Wong & Perrachione, 2007). In short, musically trained participants are particularly good listeners (Kraus & Chandrasekaran, 2010; Strait & Kraus, 2011).

With a few exceptions (Moreno et al., 2009; Thompson et al., 2004, Experiment 3), all of the studies discussed so far were correlational or quasi-experimental, which precludes inferences of causation. Although it is reasonable to assume that performing music causes an individual to improve on listening tasks, the reverse causal direction is equally plausible. For example, those with naturally poor listening
abilities (e.g., such as those measured by tests of music aptitude) would be unlikely
to pursue music training, particularly for years on end, thus guaranteeing a positive
association between listening abilities and music training, and listening and duration
of training. Conversely, musically talented individuals (with naturally good listening
abilities, or high levels of music aptitude) would be especially likely to persevere
with the demands of advanced music training. Finally, high-functioning children
(i.e., with high IQs) may perform better than other children on listening tests, and be
more likely to take music lessons.

B. Music Training and Memory

Compared with nonmusicians, musicians sometimes exhibit better memory for
auditory stimuli, including familiar music, unfamiliar music, and environmental
sounds, but not for visual stimuli (Cohen, Evans, Horowitz, & Wolfe, 2011). One
group of researchers (Chan, Ho, & Cheung, 1998) examined memory for lists
of spoken words. Women with music training had better verbal memory than
untrained women, whereas visual memory (reproducing line drawings) was equiva-
Ient across groups. A similar verbal-memory advantage for musicians has emerged
in samples of men and women (Brandler & Rammsayer, 2003; Chin & Rickard,
2010; but see Helmbold, Rammsayer, Altenmüller, 2005) or boys (Ho, Cheung, &
Chan, 2003). Indeed, musicians exhibit enhanced memory for color names they
hear but not for colors presented visually (Tierney, Bergeson, & Pisoni, 2008), and
for sequences of tones but not for sequences of letters (Williamson, Baddeley, &
Hitch, 2010). An advantage for musically trained individuals is also evident in
tasks that require auditory stimuli to be imagined rather than heard, but not in
visual-imagery tasks (Aleman, Nieuwenstein, Böcker, & de Haan, 2000). In some
instances (Franklin et al., 2008; Pallesen et al., 2010; Parbery-Clark et al., 2009;
Posedel et al., 2011) but not in others (Strait et al., 2010), music training is associ-
ated positively with performance on tests of auditory working memory.

Musically trained individuals also exhibit enhanced memory for prose, either
stories (Jakobson et al., 2003) or song lyrics that are spoken or sung (Kilgour,
Jakobson, & Cuddy, 2000). Music training may enhance the processing of temporal
order for auditory stimuli (Koh, Cuddy, & Jakobson, 2001), which could mediate
the link between training and verbal memory (Jakobson et al., 2003). Another
possibility is that musically trained individuals are higher functioning in general,
such that enhanced verbal memory is an artifact of better cognitive abilities
(Schellenberg, 2011a).

In line with this “generalist” position, performance on measures of verbal mem-
ory (for lists of words) and visual memory (for line drawings) is sometimes
enhanced among musicians compared with nonmusicians (Jakobson, Lewycky,
Kilgour, & Stoesz, 2008), a finding that conflicts with others suggesting that
memory advantages for musicians are restricted to auditory stimuli (Cohen et al.,
2011). Other researchers have also reported better auditory and visual memory
among musically trained compared with untrained children (Degé, Wehrum,
Stark, & Schwarzer, 2011). Moreover, memory for lists of numbers presented
visually is sometimes enhanced among musicians compared with nonmusicians, just as it is for sequences of tones (Greene & Samuel, 1986). In fact, in one study of older adults (Hanna-Pladdy & MacKay, 2011), music training was associated with nonverbal but not with verbal memory.

Thus, it remains an open question whether memory advantages for musicians are limited to auditory stimuli, and even whether their memory is better for auditory than for visual stimuli. Presumably, the conflicting findings depend on the particular tasks and samples. For example, when associations between music training and visual memory are not evident in samples of Chinese participants (Chan et al., 1998; Ho et al., 2003), the null result may be due to relatively good performance among all participants because of exposure to a logographic writing system. Finally, as with the listening studies, the correlational and quasi-experimental designs of the memory studies preclude inferences of causation.

C. Music Training, Vocabulary, and Reading

Music training is associated with enhanced performance on tests of language ability. In one study, 4- to 6-year-olds were assigned quasi-randomly to an intensive 4-week computer-based program in music or visual art (Moreno, Bialystok, et al., 2011). In both programs, the children were taught in groups. The music program focused primarily on listening skills and did not involve training on a musical instrument. The visual-art program was closely matched except for content. Before and after the training, both groups were administered two subtests from a Wechsler IQ test that is designed for young children. One was a measure of spatial reasoning (Block Design); the other was the Vocabulary subtest. Whereas the two groups performed similarly at pretest on both measures, only the music group showed significant improvement from pretest to posttest, but only on the test of vocabulary. The experimental design allowed the authors to infer that music-based listening training caused increases in vocabulary. Converging evidence comes from a quasi-experiment in which a year of music lessons was associated with improvements on tests of vocabulary and understanding prepositions (Piro & Ortiz, 2009). The children in Moreno et al.’s sample also showed a posttest advantage in mapping words with arbitrary visual symbols when pretest levels of performance were held constant (Moreno, Friesen, & Bialystok, 2011).

Reviews of links between music training and reading abilities have reached conflicting conclusions. In a paper that included two separate meta-analyses (Butzlaff, 2000), a survey of correlational studies led to the conclusion that music training and reading abilities are associated positively. For experimental studies, however, there was no effect, which raises doubts about any causal connection between music training and reading. A more recent meta-analysis of a larger collection of experimental studies reached a positive conclusion, but its broad definition of music training led to the inclusion of many studies with interventions that are atypical of most music lessons (Standley, 2008). We now turn to studies of associations between typical music lessons and reading abilities published since Butzlaff’s meta-analyses.
Gromko (2005) compared classes of kindergarten children who received music lessons for 4 months to other classes with no lessons. The music classes showed larger improvements in phonological awareness, which is predictive of reading outcomes, as noted in the section on music aptitude. More impressive results come from a study that randomly assigned individual kindergarteners to programs in music, phonological awareness, or sports (Degé & Schwarzer, 2011). The training involved daily 10-minute sessions for 20 weeks (100 sessions in total). At the beginning of the study, the three groups did not differ in phonological awareness, fluid intelligence, or SES. At the end, phonological awareness had improved substantially among the children who were specifically trained in these abilities, but the music group had virtually identical improvements. These findings do not appear to be the consequence of maturation because the improvements did not extend to children in the sports group.

In another experimental study (Moreno et al., 2009), 8-year-olds were assigned to one of two training programs for 6 months, either music or painting. In both groups, the children had 75-min lessons twice per week. The music training consisted primarily of listening exercises based on Kodály, Orff, and Wuytack pedagogies. A word-reading test was administered before and after the training sessions. It involved pronouncing words that varied in terms of letter-sound correspondences, with some words regular, others complex but with consistent letter-sound correspondences across words, and still others that were complex and inconsistent. Improvements from pretest to posttest were evident only for the music group tested with inconsistent words. Although this study does not allow the conclusion that music lessons improve reading per se, we can infer that listening-based music training facilitates the ability to pronounce irregularly spelled words presented in isolation. Results from quasi-experiments have revealed a similar pronunciation advantage for irregularly spelled words among adults with music training (Bugos & Mostafa, 2011; Jakobson et al., 2008; Stoesz, Jakobson, Kilgour, & Lewycky, 2007). In some studies of younger (Patston & Tippett, 2011) and older (Hanna-Pladdy & MacKay, 2011) adults, however, the effect fell short of statistical significance, or advantages in reading for musically trained children disappeared when IQ was held constant (Hille, Gust, Bitz, & Kammer, 2011). Nevertheless, on a reading task that required participants to make grammaticality judgments of individual sentences, musicians outperformed nonmusicians (Patston & Tippett, 2011).

Corrigall and Trainor (2011) measured reading comprehension among a sample of 6- to 9-year-old children with varying amounts of music training. The reading task required children to identify a missing word in a sentence or paragraph. Length of training was associated positively with reading comprehension when age and SES were held constant. The association remained evident when additional confounding variables (music aptitude, full-scale IQ [FSIQ], word-decoding abilities, or number of hours spent reading each week) were controlled, but no test was conducted with all potential confounding variables held constant simultaneously. The association was carried, however, by children who began lessons early; it disappeared when onset of training was controlled. In general, children with more training started lessons earlier than other children and they had slightly higher IQs.
Thus, evidence that music training has a direct causal effect on reading is not as strong as evidence for a causal effect between music lessons and phonological awareness or the ability to pronounce irregularly spelled words.

In a large-scale survey of more than 7,000 10th graders (Southgate & Roscigno, 2009), taking music lessons inside and outside of school made independent contributions in predicting reading achievement even after controlling for SES, race, gender, number of books read, and reading achievement in eighth grade. In a parallel survey of more than 4,000 first graders conducted by the same researchers, taking music lessons in school (but not outside of school) predicted reading achievement when SES, race, gender, number of books read, and reading achievement in kindergarten were held constant. Although these results do not inform the issue of causation, they suggest that music lessons taken outside of school are a better predictor of individual differences in reading ability for adolescents than for young children.

D. Music Training and Visuospatial Abilities

Music training is associated with enhanced visuospatial skills. In a meta-analytic review of experimental studies, Hetland (2000a) examined studies with spatial-temporal outcome measures separately from studies with other visuospatial measures. In absolute terms, the association with music training was slightly stronger in the former instance, but the median effect size was similar across the two analyses ($d \approx 0.7$). Thus, Hetland’s review provided evidence of a causal link between music training and visuospatial abilities as traditionally defined.

Converging evidence since Hetland’s (2000a) meta-analyses comes from quasi-experiments of music training and visuospatial ability that reveal enhanced performance for musicians on tests of line orientation (Sluming et al., 2002), mental rotation (Sluming, Brooks, Howard, Downes, & Roberts, 2007), and memory for line drawings (Jakobson et al., 2008). Musicians also perform better on visual-search tasks that require them to detect “embedded” figures in line drawings (Stoesz et al., 2007) or a small change in a complicated visual design (Patston & Tippett, 2011). Music training is also positively associated with adults’ performance on the Block Design subtest (a visuospatial test included in comprehensive measures of IQ) and at copying “impossible” (i.e., Escher-like) figures (Stoesz et al., 2007). When asked to discern whether a dot is flashed on one side or the other of a target line, musicians are faster than nonmusicians (Patston, Hogg, & Tippett, 2007), whether or not the line is removed before presentation of the dot (Brochard, Dufour, & Deprès, 2004). In general, musicians appear to have less of a laterality bias compared with nonmusicians (Patston, Hogg, et al., 2007; Patston, Kirk, Rolfe, Corballis, & Tippett, 2007). For example, when asked to bisect a horizontal line in two, nonmusicians tend to place their mark systematically to the left of center whereas musicians show a smaller rightward bias (Patston, Corballis, Hogg, & Tippett, 2006). The reduction in laterality may be a consequence of increased use of the nondominant hand when playing a musical instrument, which typically involves the use of both hands.
In a study of 4- to 6-year-olds, children were assigned to classes that had Kindermusik training for 30 weeks (Bilhartz, Bruhn, & Olson, 2000). There was at least one music class and one control class (i.e., no training) for each of three SES levels (low, middle, or high). Children were administered six subtests from a standardized test of intelligence before and after the lessons. On one subtest, improvements were larger for the music groups than for the control groups, and particularly large for children whose families were compliant with the instructions of the Kindermusik program. This test (Bead Memory) required children to reproduce strings of beads that varied in color and shape after viewing a picture of a target string of beads.

In another study of kindergarteners to second graders, some of the children took keyboard lessons for 8 months during one or more academic years (Rauscher & Zupan, 2000; Rauscher 2002). They were compared with children without lessons on three tests of visuospatial abilities. Children who took lessons for 3 years (kindergarten, first grade, and second grade) improved significantly on two of three tests after the first year and continued to improve (but not significantly) thereafter. Children who took lessons for 2 years (i.e., during kindergarten and second grade) showed improvements after the first year of lessons on the same two tests, a decline after the year without lessons, but additional improvement after the second year of lessons. Children who took lessons for only 1 year, in second grade, did not show any significant improvements. These results suggest that music training may have a stronger influence on visuospatial skills if it is begun early in life, a finding corroborated by correlational evidence (Costa-Giomi, Gilmour, Siddell, & Lefebvre, 2001).

After 2 years of music instruction were provided to low-SES 7- to 9-year-olds who were having difficulty in school, the children were better able to memorize and copy line drawings than were control children who did not have the intervention (Portowitz, Lichtenstein, Egorova, & Brand, 2009). Rauscher and Hinton (2011) described the results of two related unpublished studies of at-risk preschoolers enrolled in Head Start schools. In one, children were assigned randomly to 48 weeks of piano lessons, computer lessons, or no lessons over the span of 2 years. Although the groups performed similarly on a battery of 26 tests before the lessons began, the piano group performed better on visuospatial tests when the lessons were over. In a second study with the same battery of tests, children received 48 weeks of piano, voice, percussion, or no lessons. The three music groups outperformed the no-lessons group on tests of visuospatial abilities at posttest and the effects continued to be evident for 2 years after the lessons ended. It is unclear, however, whether the researchers corrected for multiple tests in either study. Moreover, because the original studies did not undergo peer review and methodological details in the review article are sparse, it is impossible to determine whether the design, procedure, and analyses were optimal. Results are similarly inconclusive from a study with no control group that reported enhanced visuospatial abilities among kindergarteners who received a year of keyboard training (Zafranas, 2004). Nevertheless, the available data point to an association between music training and visuospatial ability.
that appears to be caused by taking music lessons, with stronger effects when lessons are begun in early childhood.

### E. Music Training and Mathematics

Associations between music training and mathematical abilities are more elusive than associations with other aspects of cognition. A meta-analysis of correlational studies concluded that there is a small positive association between music training and mathematical abilities (Vaughn, 2000). When Vaughn explored the issue of causation by examining six experimental studies, she again reported a small but significant result. Three of the studies were actually quasi-experiments, however, three were unpublished, and one unpublished study was based on positive and unreliable preliminary findings that disappeared by the time the study ended (Costa-Giomi, 1999). In short, there was little evidence of a causal association between music training and mathematical skills.

Since Vaughn’s (2000) review, one quasi-experimental study examined the association between musicianship (primarily music training) and mathematical ability among high-school students (Bahr & Christensen, 2000). A significant but modest positive association was evident for some test items but not for others. Another quasi-experiment (Cheek & Smith, 1999; not included in the Vaughn meta-analysis) compared eighth graders with more than 2 years of private music lessons to a control group without lessons. The music group had higher scores on a standardized test of math ability, and students with keyboard lessons scored particularly well. In a large sample of first graders, mathematical achievement was associated with taking music in school (but not with taking lessons outside of school), even after holding constant SES, race, gender, amount read, and math abilities measured a year earlier (Southgate & Roscigno, 2009). In a study of preschoolers, early numerical concepts were better developed among children in a music program than in other children, but this association appeared to stem from differences in the home musical environment (Geoghegan & Mitchelmore, 1996; not included in the Vaughn meta-analysis). Finally, in a description of two unpublished experimental studies (Rauscher & Hinton, 2011), preschoolers and elementary-school students assigned to music lessons had larger increases on tests of arithmetic than did children in control groups.

Other researchers have failed to find an association between music training and mathematical abilities. One study compared high-school students with or without two or more music credits per year (Cox & Stephens, 2006). Grades in math courses did not differ between groups. Another study examined more than 7,000 10th graders and found that performance on a standardized test of mathematical abilities was independent of taking music lessons inside or outside of school (Southgate & Roscigno, 2009). A third study asked the question in reverse, examining whether people with mathematical training would also show higher than average musical ability. An online survey was administered to a large sample of scholars, all of whom had a doctoral degree (Haimson et al., 2011). Some were mathematicians recruited from the American Mathematical Association; others were language or linguistics scholars recruited from the Modern Languages
Association. Responses were similar between the two groups on all measures of musicianship or musicality, suggesting that training in mathematics is not associated with heightened musical abilities.

In sum, evidence that music lessons cause increases in math ability is far from conclusive. Moreover, when small associations between music training and mathematical ability are evident in correlational and quasi-experimental studies, they could be the consequence of individual differences in general intellectual ability, with high-functioning children being more likely than other children to take music lessons and to perform well on tests of mathematics.

F. Music Training and General Intelligence

The issue of whether associations between music training and cognition are general or specific to certain subsets of cognitive abilities can be addressed directly when standardized IQ tests are administered to participants who vary in music training. The most common (e.g., Wechsler) tests include multiple subtests that are combined in various ways to provide indexes of verbal ability, spatial ability, processing speed, and working memory, as well as an aggregate measure of FSIQ. These tests are standardized based on age from large samples of the general population, with good reliability and validity. Because outcome scores are measured on the same scale, direct comparisons can be made among the different subtests or indexes.

In general, individuals with music training tend to have higher FSIQs than their untrained counterparts. In one correlational study of approximately 150 10- to 11-year-olds and 150 undergraduates (Schellenberg, 2006b), duration of music lessons was positively associated with FSIQ in both samples. In the child sample, the association between duration of training and FSIQ remained evident when SES and involvement in nonmusical out-of-school activities were held constant. In the adult sample, the association between playing music and FSIQ was smaller than it was for children but it remained evident with SES controlled. In both samples, the association was strongest for aggregate measures of cognitive ability (i.e., FSIQ or the principal component extracted from the subtests), and no association between duration of training and specific cognitive abilities was evident when an aggregate measure was held constant. Interestingly, for the children, duration of training correlated significantly with performance on 11 of 12 subtests, the exception being Object Assembly, a “spatial-temporal” test that Rauscher and her colleagues believe to be linked most strongly with music training (e.g., Rauscher & Hinton, 2011).

When Corrigall and Trainor (2011) administered the WISC to 6- to 9-year-olds, the correlation between training and FSIQ \((r = .27)\) was slightly smaller in magnitude than that reported by Schellenberg (2006b, Study 1, \(r = .35\)) and only marginally significant, perhaps because of the smaller sample and the exclusion of children with no music training. In general, because a complete IQ test with all of its component subtests takes 2 hours or more to administer, these tests are rarely used in studies of music training. Instead, researchers have opted to administer briefer tests such as the Wechsler Abbreviated Scale of Intelligence (WASI) or
the Kaufman Brief Intelligence Test (K-BIT), a single measure of fluid intelligence (i.e., with no measure of crystallized intelligence or acquired knowledge), or simply one or two subtests from more comprehensive tests.

When the WASI (four subtests: two verbal, two nonverbal) was administered to 9- to 12-year-olds (Schellenberg, 2011a), musically trained children outperformed their untrained counterparts by more than 10 points (SD) in FSIQ, and the advantage for the trained group extended across the four subtests. When an even briefer version of the WASI (two subtests: one verbal, one nonverbal) was administered to 7- to 8-year-olds (Schellenberg & Mankarious, 2012), the advantage for the trained group was one standard deviation (15 points) for FSIQ and similar for the verbal and nonverbal subtests. In a comparison of musically trained and untrained adults (Gibson, Folley, & Park, 2009, Experiment 1), trained participants scored higher on the WASI by half a standard deviation. When the K-BIT was administered to adults (Schellenberg, 2011b), trained adults outperformed untrained adults again by approximately half of a standard deviation, and the advantage was evident on both verbal and nonverbal scores. Considered jointly with the results of Schellenberg (2006b), taking music lessons in childhood appears to have a stronger association with IQ when it is tested in childhood than adulthood. In any event, music training is clearly associated with general intelligence, at least as measured by the various subtests included in standard or brief IQ tests.

Claims of special associations between music training and specific aspects of cognitive ability are valid only if the association remains evident when a measure of FSIQ is held constant (Schellenberg, 2008, 2009). Indeed, controlling performance on a single subtest can lead to misleading findings (e.g., Jakobson et al., 2008; Stoesz et al., 2007). For example, Schellenberg (2009) demonstrated that the association between Block Design and music training was significant when Vocabulary but not FSIQ was held constant. Very few studies have reported an association between music training and a specific cognitive ability with FSIQ controlled. In one exception (Corrigall & Trainor, 2011), the partial association between reading comprehension and duration of training remained evident when FSIQ was held constant. Because all of the children in their sample had some music training, however, the association between duration of training and FSIQ was itself weak and short of statistical significance.

The simplest interpretation of results from correlational studies and quasi-experiments is that high-functioning children are more likely than other children to take music lessons and to perform well on most tests they take. Nevertheless, Schellenberg (2004) provided evidence that the causal direction may also go from music training to cognitive abilities. He assigned 144 6-year-olds randomly to a year of music lessons (keyboard lessons or Kodály classes), drama lessons, or no lessons. The lessons were conducted in groups of six children. All of the teachers had similar qualifications and all of the classes were taught in the same location. All children were tested with the entire WISC before and after the lessons. The 12 children who dropped out before the posttesting session did not differ from other children in terms of FSIQ at pretest. Increases in FSIQ were about three points larger, however, for the two music groups combined (who did not differ) than for
the two control groups combined (who did not differ), and the advantage extended across the subtests and indexes. Only the drama group had significant improvements in adaptive social skills as measured by parent reports. Random assignment to the different conditions allows us to infer that music lessons caused small increases in cognitive abilities and that drama lessons caused improvements in social skills. At the same time, the findings do not preclude the possibility that high-functioning children are more likely than other children to take music lessons.

In the study by Moreno et al. (2009), 8-year-old children were administered the WISC before and after 6 months of training in music or painting. Increases in FSIQ from pretest to posttest were 5 points greater in the music group than in the painting group, a larger effect than that reported by Schellenberg (2004). Because of unusually large test/retest improvements (i.e., 12 points for the music group, 7 for the painting group) due to the short time frame of the study, the difference between groups was not significant. In general, consecutive administrations of the same IQ test should be separated by a year or more.

G. Music Training and Academic Achievement

Associations between music training and general cognitive ability extend beyond intelligence testing to grades in school. For example, Wetter, Koerner, and Schwaninger (2009) examined the academic records of third- to sixth-grade Swiss children who either did or did not take music lessons outside of school. Children with lessons had higher average grades even when SES was held constant, and the advantage was evident across all school subjects except for sports.

Similarly general associations are evident in samples of students from the US. Fitzpatrick (2006) examined performance on standardized tests of academic proficiency for more than 15,000 American high-school students, more than 900 of whom were registered in an instrumental music course. The music and control groups were further subdivided into low- and normal-SES groups. Fitzpatrick looked at performance in fourth, sixth, and ninth grades, before the students opted to take a music course in high school. With differences in SES held constant, the future instrumentalists outperformed the control group in every subject at each grade level. These results confirm that high-functioning children are more likely than other children to take music lessons, at least in high school. Other evidence implies that (1) positive associations between taking music classes in school and performance on standardized tests are more likely when the instruction is of particularly high quality (Johnson & Memmott, 2006), and (2) replacing some standard academic classes with instrumental music lessons does not have a negative impact on average grades in elementary school (Kvet, 1985).

Gouzouasis, Guhn, and Kishor (2007) examined performance on standardized tests of academic achievement among 150,000 Canadian 12th-grade students. The tests provided separate scores for mathematics, English, and biology. Compared with other students, those who took music classes in 11th grade had higher scores in mathematics and biology but not in English. Among the music students, grades in 11th-year music courses were correlated positively with 12th-year standardized
scores for mathematics and biology, and weaker but still evident for English. By contrast, participation in 11th-grade visual arts courses was not associated with standardized test scores in 12th grade. These results confirm that high-functioning students are more likely than other students to take high-school courses in music but not in visual arts, that taking music courses in high school does not interfere with achievement in core academic subjects, and that students who do well in music classes tend to do well in other subjects.

Similar findings in Schellenberg’s (2006b) study revealed that among elementary-school children, as duration of training increased, so did academic performance whether it was measured with actual grades on report cards or performance on a standardized test, and even when SES and duration of nonmusical activities were held constant. Even more provocative was the finding that school performance was associated positively with duration of training when IQ was controlled. In other words, children who took music lessons for years on end tended to be particularly good students. Duration of playing music regularly in childhood also predicted high-school average when SES was controlled. Finally, there is experimental evidence that 1 year of music training causes small improvements in academic achievement (Schellenberg, 2004).

The association between duration of music training and general cognitive abilities suggests that professional musicians should be geniuses. When highly trained individuals or professional musicians are compared with similarly professional individuals without music training, however, the association often disappears. In other words, although professional musicians may be above average in intelligence, people who are highly trained in other disciplines perform at similar levels. For example, when members of a symphony orchestra or students from university music departments were compared with students from other disciplines (e.g., psychology, business) with a similar amount of education, the IQ advantage for the music students vanished (Franklin et al., 2008; Helmbold et al., 2005) or favored the students from the nonmusic disciplines (Brandler & Rammsayer, 2003). Similarly, when the comparison involved highly trained versus untrained participants, differences in intelligence fell short of statistical significance (Sluming et al., 2002) or disappeared altogether (Bialystok & DePape, 2009; Patston & Tippett, 2011; Schellenberg & Moreno, 2010). Thus, it appears that music training is associated positively with intelligence when training is added as an activity on top of regular schooling.

One problem with this interpretation is that many of the null findings involved tests of fluid intelligence (Bialystok & DePape, 2009; Brandler & Rammsayer, 2003; Franklin et al., 2008; Helmbold et al., 2005; Patston, Hogg, et al., 2007; Patston & Tippett, 2011; Schellenberg & Moreno, 2010)—such as the Cattell Culture Fair Test, Raven’s Progressive Matrices, or the Matrices subtest from the Wechsler tests—rather than more comprehensive measures that include subtests of learned abilities (e.g., vocabulary). Although the null findings might therefore be interpreted as showing that associations between music training and cognitive abilities do not extend to pure measures of fluid intelligence, other researchers have reported such an association (Degé, Kubicek, & Schwarzer, 2011; Forgeard, Winner, et al., 2008; Hille et al., 2011; Portowitz et al., 2009; Thompson et al.,
In short, the distinction between music training in childhood and whether one is a “real musician” appears to be an important one, such that associations with general cognitive ability are much more likely in the former case.

**H. Music Training, Social-Emotional Abilities, and Executive Functions**

Are nonmusical associations with music lessons strictly cognitive? The answer appears to be yes. In a correlational study of 6- to 11-year-olds, duration of music training was independent of social skills as measured by parent reports (Schellenberg, 2006b). In a 3-year experimental study with random assignment of low-SES fourth-graders to music lessons or no lessons, the two groups did not differ in self-esteem at the beginning of the study or at the end of each of the three years (Costa-Giomi, 2004). In a 1-year experimental study, music training was not associated with improvements in social skills (Schellenberg, 2004). In a quasi-experimental study, musically trained and untrained undergraduates performed similarly on a test of emotional intelligence even though they differed markedly in FSIQ (Schellenberg, 2011b). In another quasi-experimental study of 7- and 8-year-olds (Schellenberg & Mankarious, 2012), musically trained children outperformed untrained children on a test of emotion comprehension but the advantage disappeared when FSIQ was held constant. Most of the null findings involved music lessons taught individually, however, and there is some evidence that relatively intensive group music interventions may promote social development (Kirschner & Tomasello, 2010; Rabinowitch, Cross, & Burnard, 2012), even among infants (Gerry, Unrau, & Trainor, 2012). Results from more standard music classes taught in schools are equivocal (Rickard, Appelman, et al., 2012; Rickard, Bambrick, & Gill, 2012).

Some researchers have attempted to identify the mechanisms that drive the association between music training and general cognitive abilities. Although mental or perceptual speed may play a role (e.g., Bugos & Mostafa, 2011; Gruhn, 2006), there has been more speculation that executive functions act as mediating variables (e.g., Hannon & Trainor, 2007; Jäncke, 2009; Schellenberg & Peretz, 2008). Executive functions refer to a set of mechanisms that are involved in conscious control of thought, including working memory, inhibiting inappropriate responses, planning ahead, flexibility, concentration, selective attention and ignoring irrelevant information, the ability to change strategies as the situation demands, and so on. Executive functions are correlated with IQ (Salthouse, 2005; Salthouse, Atkinson, & Berish, 2003) and amenable to influences of training, especially in childhood (Dowsett & Livesey, 2000; Kloo & Perner, 2003; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). The hypothesis is that music training improves executive functions, which in turn lead to better performance on a wide variety of cognitive tests. In line with this mediation hypothesis, children who take piano lessons believe that the lessons help them to develop skills of concentration and discipline (Duke, Flowers, & Wolfe, 1997). To date, however, objective tests of the hypothesis are inconclusive.
For example, in one study of musically trained and untrained adults matched for fluid intelligence, trained individuals were faster on tests that required them to ignore conflicting information, such as identifying whether a sung pitch was high or low while ignoring whether it was sung with the word high or low, or identifying whether an arrow pointed left or right while ignoring whether it was on the left or the right side of a display (Bialystok & DePape, 2009). Although the trained group was faster than the untrained group in the conflicting conditions on both tests, the groups were similarly disadvantaged in the conflicting compared with the consistent conditions. In another quasi-experimental study, musically trained adults scored higher than their untrained counterparts on a test of IQ and on tests of executive function that measured verbal fluency and divergent thinking, but the researchers did not test whether the association with FSIQ was mediated by executive function (Gibson et al., 2009). Musicians have also outperformed nonmusicians on a go/no-go task that required them to respond quickly to a beep but not to a siren, a task that measures attention and response inhibition (Strait et al., 2010).

In one experimental study, older adults were provided with 6 months of piano lessons or no lessons (Bugos, Perlstein, McCrae, Brophy, & Bedenbaugh, 2007). The music group but not the control group improved over time on a subtest from the WAIS (Digit-Symbol Coding), but the test measured processing speed rather than an executive function. Another test (Trail Making) measured one aspect of executive function (i.e., attention control). In a baseline condition, participants connected dots on a paper sequentially according to number (1, 2, 3...). In a subsequent test condition, participants connected dots sequentially according to number and letter (1, A, 2, B, 3, C...). Only the music group improved over time in the test condition, but the effect disappeared when performance in the baseline condition was held constant. Moreover, in a quasi-experiment that measured performance on the Trail Making task with baseline performance controlled, music majors performed similarly to undergraduates in other disciplines (Isaacs & Trofimovich, 2011). Finally, in a quasi-experimental study of older adults (Hanna-Pladdy & MacKay, 2011), music training was associated with faster performance in both the baseline and test conditions of the same task (see also Bugos & Mostafa, 2011), but the difference between conditions was similar for trained and untrained participants.

Studies of children are similarly inconclusive. In one quasi-experimental study of 9- to 12-year-olds (Schellenberg, 2011a), musically trained children had substantially higher FSIQs than untrained children but the groups performed similarly on four of five tests of executive function (verbal fluency, tower of Hanoi, Wisconsin Card Sort Test, and sun-moon Stroop). Although the music group was better on the Digit Span test of working memory, this test is also included in the Wechsler measures of IQ, which highlights the rather loose distinction between intelligence and measures of executive function (Salthouse, 2005). In a sample of children just slightly older, however, musically trained children performed better than untrained children on a test of verbal fluency (Hassler, Birbaumer, & Feil, 1985). In another sample of 9- to 12-year-olds (Degé, Kubicek, et al., 2011), musically trained children performed better than untrained children on a test of fluid intelligence and on five different measures of executive function. Moreover, the association between
training and intelligence disappeared when measures of selective attention and inhibition were held constant, providing evidence for the mediation hypothesis described above. Finally, when 4- to 6-year-olds were assigned to 4 weeks of computer-based music or visual-art training, only the music group had significant improvements on a go/no-go task that required them to press a button when a white but not a purple shape was presented (Moreno, Bialystok et al., 2011). In conclusion, because the results differed markedly across studies, evidence for the mediation hypothesis is equivocal.

It is also important to clarify that some researchers have failed to find an association between music training and general cognitive abilities. For example, in one study, cognitive abilities (measured by standardized tests and by academic achievement) of low-SES fourth graders at the end of a 3-year intervention were similar between the piano and control groups (Costa-Giomi, 1999, 2004). In another study, when music lessons took the place of mathematics or language classes beginning at the sixth-grade level, the effect on cognitive performance was negligible even after 3 years of the intervention (Zulauf, 1993/1994). Moreover, students who register in music courses in school sometimes have average grades or IQ scores similar to those of students who do not take music courses (Cox & Stephens, 2006; Degé, Kubicek, et al., 2011). Finally, as noted earlier, the association between training and cognitive abilities often breaks down when “real musicians” are compared with nonmusicians.

VI. Conclusions

Music aptitude is associated with linguistic abilities, including phonological processing, facility with acquiring a second language, and in some instances, reading, whereas the notion of a special link between natural musical and mathematical abilities has virtually no empirical support. Moreover, many of the positive findings may be attributable to a more general association between music aptitude and cognitive functioning. Indeed, associations between music aptitude and general cognitive abilities, including performance in school, are often strong, particularly in childhood. Notable exceptions involve cases of musical savants as well as individuals with amusia.

Studies of cognitive performance after listening to music do not support the proposal of a special link between listening to music composed by Mozart and visuospatial abilities. Rather, the effect is a consequence of music’s ability to improve the arousal level and mood of the listener, which, when elevated, improve many aspects of cognitive processing. Temporary changes in arousal or mood caused by music listening can have a range of cognitive benefits, from improving creative drawing in children to visuospatial performance in adults. The available evidence also indicates that any music favored by the listener can temporarily improve arousal or mood and elevate cognitive performance.

Studies of the effects of background music on cognitive abilities have reported many contradictory findings. The inconsistencies appear to be due to a number of
factors, including the difficulty of the task and the amount of working memory required, the modality of the task (i.e., visual versus auditory), the individual’s personality, and specific attributes of the music. In short, background music can cause improvements in cognitive performance (e.g., better math abilities among children with behavioral problems) as well as decrements (e.g., poorer reading comprehension among adults who hear music that is both loud and fast). Underlying mechanisms that are likely to contribute to inconsistent findings include the listener’s emotional response to the music and cognitive interference.

Music training is associated with enhanced performance on a wide variety of listening tasks, musical or otherwise. Music training in childhood also tends to be a predictor of good performance across a wide variety of cognitive tests, including tests of memory, language, and visuospatial abilities. Music training is also associated positively with general intelligence and school performance. By contrast, comparisons of adult musicians and nonmusicians often yield null findings when the outcome measures do not involve music or listening. Regardless, the correlational and quasi-experimental designs that typify the vast majority of the available research preclude clear inferences of causation whatever the outcome variable. The available evidence suggests that high-functioning children (i.e., higher IQ, better performance in school) are more likely than other children to take music lessons and to perform well on a variety of tests of cognitive ability, and that music lessons exaggerate these individual differences slightly. Cognitive advantages for individuals who become musicians in adulthood are less consistent except on listening tasks.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada. We thank Ellen Winner and Swathi Swaminathan for their helpful comments on an earlier draft.

References


perspectives on judgment and decision research (pp. 365–393). New York, NY: Cambridge University Press.


