THE UNIVERSITY OF HULL

Working memory for music, pitch labels and solfège:

A cross-cultural study of university students' aural and cognitive skills

being a Thesis submitted for the degree of PhD

in the University of Hull

by

Artemis Apostolaki

June 2012

Contents

Acknowledgements

Introduction

The largest part of the extant research on music memorisation (Samson & Zatorre, 1991; Chaffin & Imreh, 1997; Williamon, 1999; Ginsborg, 2002) identifies four basic types of music memory: aural, visual, kinaesthetic and conceptual, each one of which has been studied extensively by psychologists and musicians in order to help musicians understand better the way they learn and memorise musical material. None of these methods, however, accounts for the memorisation of note sequences as verbal material, practised mostly by musicians who use the music-reading method known as solfeggio, or solfège. The initiative for the present study was born after a related incident during my Masters degree in the UK, when I was asked to do a last-minute substitution for a xylophone player in a wind band concert. As both percussionists had cancelled, the xylophone-and-marimba solo in one of the pieces had to be performed by me – a cellist with no previous experience in percussion – and a viola player, who was also a self-taught percussionist. The solo itself was not difficult: it lasted no more than 10 seconds and both instruments were playing in unison; at the end of the second, and last, rehearsal, I found myself playing it by memory in order to be able to concentrate on hitting the correct notes and occasionally look at the conductor. The marimba player, on the other hand, had chosen to look at the music and the instrument, omitting the eye-contact with the conductor, resulting in some criticism from the afflicted party and a lively discussion after the rehearsal regarding our memorisation skills. I had learnt the solo as a nonsense poem consisting of sequences of solfège syllables; the other percussionist explained he would not feel safe performing by memory on a relatively unknown instrument to him, especially at such short notice. This led to the realisation that two different kinds of memorisation were being discussed: the marimba player was clearly talking about an insecurity stemming from the lack of kinaesthetic cues; but which memorisation strategy did solfège syllables represent? The verbal, or quasi-verbal, element pointed to conceptual memory; on the other hand, this type of memory is primarily linked to meta-musical information about structure and patterns, of which I was only indirectly aware, through the knowledge of the pitch material. This incident raised numerous questions: do people who know solfège use it as a separate strategy for

memorising music? Do people who do not use solfège have a similar way to approach music, through which they can be constantly aware of the pitch material in a piece or passage? Does the use of solfège consequently create an advantage in the music memorisation process? Do solfège syllables constitute language, music, or both?

In order to investigate these questions, some definitions have to be given, starting with the definition of 'solfège' itself. According to *Grove Music Online*, solfège is

> The use of syllables in association with pitches as a mnemonic device for indicating melodic intervals. Such syllables are, musically speaking, arbitrary in their selection, but are put into a conventionalized order…
> (Hughes and Gerson-Kiwi, *Grove Music Online*)

It was devised most probably between 1025 and 1030 by Guido d' Arezzo, an Italian Benedictine monk, well known for his pioneering work on music methodology: the arrangement of staff notation in thirds as well as the invention of a set of overlapping kinaesthetic-based systems, part of which is solfège, are attributed to Guido (Miller, 1973; Hughes and Gerson-Kiwi, *Grove Music Online*). Guido's innovative ideas stemmed from the desire to make the teaching of church music easier for both students and instructors: solfège, thus, was initially conceived to serve as a memory aid and a method to sight-read and learn new music. Up to then, songs were mostly taught by rote and continuous repetition, a method which was extremely time-consuming and counterproductive; Guido decided to turn his attention from song-learning to *music* learning in general, so that choir singers would be able, first of all, to *hear* consecutive pitches accurately and then also be able to sing them. In order to achieve this, he attributed a different syllable to each pitch, so the piece could also be taught as a series of syllables; there are various theories regarding the origins of the first solfège syllables (ut, re, mi, fa, sol, la and arguably si), the dominant ones being either that Guido adopted the syllables of Arabic solfège or that he used the beginnings of verses from a hymn to St. John (Miller, 1973). What is important is that the system Guido devised – or, at least, perfected – proved so successful that it gained the approval of Pope John XIX; having the blessing of the head of the Catholic church, solfège spread easily in Europe and its colonies over the following centuries.

During the millennium of solfège history, the uses, connotations and basic meaning of the term have fluctuated drastically. The word is frequently met in its French, Italian and several English variants: 'solfège', 'solfeggio', 'solmization' or 'sol-fa' respectively; all the above are found almost equally often in the literature, either used interchangeably or having subtle differences in focus, according to the different school of teaching they refer to. For example, the word 'solmization' usually implies a stress on the aural and mnemonic aspect of the system, while 'solfeggio' or 'solfège', usually employed as synonyms, often suggest a link with the Italian vocal technique and the tradition of the Conservatoire de Paris. Hereon, the term 'solfège' is going to be used in all cases, except when quoting directly from a source which uses another version of the word. The decision to use the term 'solfège' consistently in the present study was made mainly due to the existence of the word in the title of one of the most essential solfège methods, first published in France at the end of the 19th century (Dannhäuser, Lavignac and Lemoine, 1891): 'Solfège des Solfèges' is an influential textbook, which has dominated the field since its first publication and has been used in almost every country that incorporates solfège in its music educational system. The word 'solfège' in the present text will denote **the practice of assigning the verbal labels 'do', 're', 'mi', 'fa', 'sol', 'la', 'si/ti', to particular acoustic frequencies within a musical context**.

Another important distinction has to be drawn between fixed-do and moveable-do systems: in fixed-do solfège, every syllable always corresponds to a specific pitch; for example 'do' is always C, 're' is always D and so on. Moveable-do, or tonic sol-fa, syllables, on the other hand, denote degrees of the scale rather than specific pitches; in this instance, 'do' is always the tonic. The moveable do system has further subdivisions, such as the la-minor system, in which 'la' (or A) is always the tonic of minor scales; in addition, most moveable do systems employ different vowels for accidentals: sharpened notes change their vowel to [i] and flattened notes to [e], for example la (A) becomes li (A#) and le (A♭) respectively. The issue of whether fixed- or moveable-do is preferable has been the subject of a long debate with strong arguments on both sides (Siler, 1956; Bentley, 1959; Humphreys, 2006); the superiority of the fixed-do system in terms of perceptual refinement and appropriateness are going to be argued in Chapter 1.2 (Solfège as perception). According to the definition of solfège provided in the previous

paragraph, whenever the term 'solfège' is encountered in the present review, it is meant to denote the fixed-do system, unless otherwise indicated.

Institutions for Western music around the world have all used, at one time or another, one or more of the aforementioned solfège systems to teach music; the contemporary music scene can be divided in two in respect to solfège teaching in the music curriculum:

1. Countries in which solfège is a fundamental part of music education, such as Russia, France, Spain, Italy, Balkan countries, Latin American countries and others.

2. Countries in which solfège is only taught by some higher education institutions for music, depending mostly on each institution's choice of curriculum, staff and approach; such countries are the USA, the UK, Germany, Australia and Canada.

As the present research aims to look at solfège in relation to music memorisation, there is a variety of issues that need to be discussed: solfège, aural skills and training, music cognition and perception, memory and working memory, music memory and memorisation are some of them. As each one is broad enough an area of research to cover a book by itself, a common point of reference is necessary in order to keep the discussion focussed and comprehensible: in the following chapter, all topics are going to be seen in relation to, or in combination with, solfège.

The first chapter of the thesis is going to review the relevant literature in the field, in order to provide a theoretical grounding to the empirical investigation of the role of solfège in specific aspects of musical reality. Chapter 2 is going to discuss the methodology adopted in the empirical part of the study and establish the aims of the research. In Chapters 3, 4 and 5 a full description and analysis of the data gathered through experimental work will be provided; Chapter 6 will provide a concluding discussion of the experimental findings and theoretical arguments.

Chapter 1

Solfège in context

1.0. Introduction

The first section of the present chapter is going to present historical aspects of solfège, describing certain phases of its evolution and development in selected countries and how it has functioned as a problem or as a solution in respect to particular musical, methodological and pedagogical issues. In the subsequent section, the perceptual basis of solfège is going to be argued; finally, specialised issues of human memory and music memorisation are going to be tackled in relation to solfège.

1.1.0. Historical Perspectives

Apart from the confusion caused by the variants of the word solfège in other languages, the content of the term itself seems to be subject to alterations according to different periods' attitudes and approaches to musicianship. Even during the same time period in different countries, solfège has been regarded as a given for elementary music learning or as bold innovation; as an obsolete teaching method or as the source of inspiration for creative music learning; as a major hindrance for musical creativity or as the ideal framework for musical creation. It is out of the scope of the present study to provide an exhaustive account of global attitudes towards solfège; it is, however, interesting and worthwhile to briefly look at how certain trends in education have affected the perception and use of solfège by music teachers in countries that have had played an influential role on music learning traditions.

As research on solfège per se is scarce, the present account is based on numerous sources giving more or less detailed accounts of perspectives and attitudes towards solfège, which are usually expressed by the way of a debate on music education in a particular country, at a particular time. There is certainly more than one way in which this material can be categorised: according to geographical area or time period, according to the focus of the author – whether it is a philosophical, educational, performance-based approach or other – or according to the attitude adopted towards solfège in particular. What is more important than the order of presentation is to acquire an accurate impression of the evolution of tendencies, trends and musicians' perceptions

regarding solfège, and how, in certain countries, it came to be rejected after decades or even centuries of practice.

The historical account of the evolution of solfège will be divided in two parts: in the first, a brief history of solfège in selected countries will be presented, along with prominent figures and institutions that have supported solfège in these countries; the second part will outline the general, diachronic arguments of educators and researchers against solfège and how they are refuted in the literature.

1.1.1. A historical geography of solfège

A report of the historical course of solfège internationally is beyond the scope of the present thesis; this chapter will present certain representative aspects of solfège evolution in selected countries, in order to provide an understanding of the basic trends in music education with which solfège has been linked during the centuries of its existence.

1.1.1.1. Solfège in Continental Europe

Since its invention in early 11th-century Italy, solfège has been taught in all Italian institutions for music, ecclesiastical and, later, secular: from the medieval boys' choir in St. Mark's cathedral to the specialized 20th-century institutions for music teachers, solfège has been invariably considered an absolute necessity to the performing musician, the composer and the music teacher (Arnold, 1962-1963; Mangino, 1963; Selfridge-Field, 1976). The historical coincidence of the genesis and establishment of solfège on the one hand and the subsequent flourish and dominance of Italian music on the other should not be considered to be fortuitous: solfège, along with the development of notation, provided an appropriate background for detailed, rigorous practice as well as a novel way to approach vocal music; this, in turn, created the optimal environment for the emergence of the Italian singing school, distinguished, amongst other things, for its emphasis on technique. The spread and influence of Italian music during the

Renaissance familiarised the rest of Europe with solfège, which is presumably how it came to be standard practice in the majority of European countries; the assimilation and development of solfège in the music culture of some of these countries will be depicted in the following paragraphs.

The word 'solfège' appeared in French for the first time in 1769, along with the Italian instruction methods for singing, in an effort to raise vocal performance standards in France (Phillips, 1997); solfège training has become an integral part of music education in France ever since, having been preserved through wartime and social, educational and political reformations (Coeuroy and Baker, 1927; Wayne, 1945). One of the oldest and most prestigious institutions for music in continental Europe, the Conservatoire de Paris, has included solfège teaching in its curriculum since its establishment in late 18th century; solfège classes were not regarded as rudimentary, but instead trained students to use solfège as means to achieving excellence in the study of harmony, theory and rhythm (Philipp and Martens, 1920). Pierre Baillot, a prominent concert soloist and violin teacher in the Paris Conservatoire, co-author, along with Kreutzer, of the official Conservatoire violin method, argues that 'To have the student undertake the study of the violin before he has learned solfège is to condemn him to reading music without understanding it.' (Baillot, 1834 in Hiatt and Cross, 2006, p. 49). Of the many famous students of the Paris Conservatoire, Bizet and Debussy had received numerous prizes in solfège as children (Curtiss, 1952; Jean-Aubry, 1918; Prod'homme, 1918); Debussy had apparently studied solfège under Lavignac, one of the authors of the seminal *Solfège des Solfèges*. Another leading figure in music pedagogy, composition and conducting, who believed firmly in the necessity of solfège was Nadia Boulanger: having experienced the Conservatoire both as a student and as a teacher, she considered solfège to be of primary importance for all musicians and for composers in particular, as it helped the development of aural and rhythmic skills and forms a solid basis for creation (Ohanian, 1977). Boulanger herself had been reported to have excellent solfège skills and to use them not only for reproduction and practice, but also for improvising (Peles, 1994).

In the rest of Europe, mentions of solfège as an integral part of the national music curriculum reveal its established status as a valuable educational tool in Portugal (Kosloff, 1963), Poland (Hlawiczka, 1931), Hungary (Jardanyi, 1957; Ádám, 1965), and Bulgaria (Wassell, 1966).

1.1.1.2. Solfège in England

The status of solfège in England has been hardly as prestigious as in the rest of Europe, despite the fact that it was known to the English as early as late 16th century: in the 'Taming of the Shrew', Shakespearean characters make flawless use of solfège terminology (Long, 1950; von Ende, 1965), a fact that presumably reveals Shakespeare's own familiarity with the solfège idiom. The teaching of solfège, however, did not become instantly and consistently popular in England: performers and music teachers appear diachronically more ambivalent regarding solfège's necessity and importance, resulting in a rather heterogeneous practice of solfège in England through history. Sands (1942) reports that all 18th-century singing teachers in England supported solfège as an important technique, but not always explicitly as to the reason for this importance, i.e. whether it benefited articulation, embellishments or sight-singing; on the other hand, certain celebrated teachers at the time seem to 'lament the neglect of sol-fa in England', while others believed in its introduction after the singer had been learning intonation by intervals (Sands, 1943-1944, p. 22). A possible reason for the intentional neglect of solfège, especially in the first years of training, seem to be the complexity caused by the use of the moveable do system: if this is considered within the context of 18th century music education, which was, naturally, quite different to the contemporary specialised training musicians receive in academies and universities, it can be assumed that the demands placed by the moveable do system on the average beginning singer were excessive for the period's standards and needs. Despite all this, however, it seems that rudiments of solfège were taught by almost all teachers at the time (Sands, 1942, 1943-1944).

A century later, a letter to the editor of *Musical Times* reveals that solfège classes had become less popular, apparently due to a misconception that girls should not receive singing lessons before they are 17-18 years old; the author regrets the fact that England does not follow the example of continental Conservatoires where all music students, instead of singers only, attend solfège lessons from an early age, thus having the opportunity to become fluent sight-readers (Lutgen, 1889). Such appeals, though, do not seem to have altered drastically solfège's position in English music education: the attitude towards solfège in pre-war Britain is conveyed in a 1931 article, which describes a state of general incomprehension and distrust towards solfège, nevertheless

taught in music schools at that time (McNaught, 1931); the author himself appears to be a supporter of the moveable do method, although not as opposed to fixed do but rather as opposed to a lack of solfège training in general. Solfège's standing within the musical elite of the time appears grim: it was taught to most, understood by few and put into practice by even fewer; this continued to be the case in the next decades, resulting in the gradual restriction of solfège to elementary school classroom teaching and the subsequent disuse by most practising musicians.

1.1.1.3. Solfège in America

In the southern part of the American continent, Mexico, Central America and the rest of Latin America, solfège has been and still is an integral part of music education (Seeger, 1944; Proudfit, 1966; Inglefield, 1962; Lawler, 1947; Beattie and Curtis, 1941; Sigmund, 1972; d'Arcangues, 2004); this is presumably carrying on the traditions established by conquistadors and all later colonizers. This, however, is not the case in the north of the continent: musicians in the USA are not systematically trained in solfège from an early age and whether they learn it in higher education is left to the discretion of the staff of each music department. The same is true in Canada, although, most probably owing to the ties of the country with France, solfège is somewhat more familiar to Canadians than to American musicians; nevertheless, neither of these countries can be described as providing consistent solfège training throughout their music education systems.

Despite the fact, however, that solfège is definitely not a current priority for music education in the USA and Canada, several prominent institutions in both countries have used it extensively in the past: such is the example of the New York National Conservatory, which in 1887 included three solfège teachers in a 12-member staff (Rubin, 1990); the Oberlin Conservatory, which has included advanced solfège courses in its curriculum since the early 1960s (Gehrkens, 1960) and also the Montreal Conservatoire in Canada, whose director in the late 1920s, Eugène Lapierre, argues that the compulsory teaching of solfège in schools by competent teachers should be a priority in order to help Canadian music evolve and achieve world-class standards (Baillot, 1932 in Calvocoressi, 1932, p. 616). Other conservatories, preparatory

institutions for students or teachers and universities for the minorities across America are also reported to offer solfège classes, starting from 1799, when the first American conservatory is said to be founded in Boston by the Italian Filippo Traetta (Marraro, 1946), until about the 1950s; in the majority of these cases solfège is reported to be taught primarily as an aid to sight-singing (Schaaf, 1985; Forssmark, 1941; Peithman, 1945; Newton, 1953; Troth, 1961; Papke, 1970; Coffman, 1987; Buckner, 1982; Lowens and Lombard, 1985).

In post-war North America, however, a number of American music educators were hostile, amongst other things, towards the idea of solfège teaching in schools and in university classrooms: this hostility needs to be understood in the light of a general demand for the reformation of a problematic music educational system. In this spirit, some propose the condensation of solfège, aural training and harmony (O'Neill, 1948), or the rapid progression from, or overall rejection of, solfège singing (Benfield, 1944; Blethen, 1953), so that classroom teaching keeps students interested and motivated whilst providing at least the basic music education, despite time restrictions; overall rejection of solfège singing is also proposed by educators who favour alternative methods, such as letter names or numbers: Nye (1948) argues that numbers are much easier to learn and promote efficient sight-singing, although he admits solfège to be an optimal instrument for cultivating the voice (Nye, 1948). Rolland (1947) argues that solfège as a basis for strings teaching is technically illogical as it does not promote learning of unified fingering patterns across strings; others argue against the excessive specialisation of the undergraduate university curriculum, and, in line with this, oppose the breaking down of harmony teaching in solfège, dictation, sight-singing, keyboard harmony and so on (Wilson, 1946). The view that solfège is not an omissible sub-component of an overall more sophisticated study of music, but in itself a valuable tool for music learning and music making will be argued later on in this thesis; this argument, however, did not prevent American music teachers to keep voicing objections against the perseverance on solfège teaching in schools.

All the above led to the gradual abandonment of solfège in America during the 1950s in favour of a system that would be, or at least seem to be, more liberal and 'fun'-oriented; this happened much to the dismay of many music educators, who considered this turn to be equivalent to leaving children in elementary school musically illiterate (Nomme,

1954). In spite of the retraction of solfège from formal music education, occasional expressions of appreciation towards the spectacular results achieved in certain European and Soviet institutions and ensembles, who expressly studied solfège rigorously, can be found in the literature (Wassell, 1966; Lowe and Pryor, 1959; Lawler, 1961; Revelli, 1961); such expressions perhaps indicate that some music teachers still believed that solfège can promote musicianship, although they did not know themselves the way to handle solfège teaching and practice in a way that would be both effective and pedagogically appropriate. This awkwardness is illustrated in the observations of a questionnaire-based survey in a sample of 100 students and 14 music teachers in Pennsylvania, showing that solfège singing skills ranked quite low in both students' and teachers' preferences (Kotts-Murphy and Brown, 1986); the issue of time, place and case-appropriate incorporation of any music teaching method in a curriculum will be discussed in the following section.

1.1.2. Theoretical debate

In the previous sections, literature on solfège was presented within a historical-geographical framework. Certain writings, however, which were not case-specific or country-specific, could not be represented in this context; this section includes reviews on papers that argue against or in favour of solfège on a more theoretical basis. The basic patterns of arguments against solfège lie very closely to each other and move along two basic lines:

1. Solfège teaching is susceptible to becoming an end in itself rather than means to achieving musicianship.
2. Solfège learning is extremely demanding, which makes it a drawback for both weaker students, who find it almost impossible, and for stronger students, who waste their time on solfège instead of practising more crucial aspects of music.

The same arguments, with minor adjustments, are presented and elaborated on in Elliott (1982), who discusses the debate between advocates of music reading teaching in American public schools and those of their opponents. Music reading is, of course, distinct from the practise of solfège; on the other hand, both elements are very often

found together in approaches to music education which promote musical 'literacy' as the cornerstone of a deeper and complete musical understanding. It is not within the purposes of the present review to evaluate the pedagogical and musical value of any particular approach to music teaching; the importance of the overall framework, however, in which music learning takes place, is frequently stressed by researchers (Swanwick 1988, 1999).

1.1.2.1. Solfège vs. aesthetics and meaning

The use of solfège in music instruction is very often regarded as being opposed to the aesthetic approach to music: this is based on the common misconception that a student who learns music with solfège is likely to know more about rules and techniques and treat music as an elaborate combination of notes than to be able to comprehend and appreciate music as an art form (Coeuroy and Baker, 1929). Others support that solfège is merely a part of training towards the development of a specific skill rather than music education per se (Beardsley, 1971); in the same line of argument, the use of solfège is recommended by some authors only in cases when a specific problem, rhythmic or melodic, arises in a passage, while on all other occasions solfège is to be treated as an artificial and undesirable dissection of music (Hoskins, 1959).

The mistake all the above authors seem to be making is that they attribute to solfège properties that it cannot hold by itself: solfège functions only as a part, albeit an extremely essential one, of a general approach to music teaching, which has to be time, place and target-appropriate. If this framework is selected carefully and applied efficiently, solfège can only prove beneficial to the developing musician; consequently, potential problems do not arise simply from knowing and practising solfège, but rather from having been taught solfège without at the same time being instructed how to manage this knowledge towards the end of musicianship. This view is reflected in Leonhard (1982), who describes how solfège can function as a 'false haven' (Leonhard, 1982, p. 24) for music teachers, who use it at the early stages of music instruction simply because they do not wish to bother with matters of musical expressivity or aesthetics; this is a common fate of even the most brilliant concepts, be it solfège, the theory of relativity or the internet, when they are subject to misuse and subsequently

reduced to serving other needs than they ought to. The fact that understanding music should be placed above everything else in music teaching is supported by many music educators (Lamers, 1960; Castaldo, 1969): Castaldo (1969) even proposes that, apart from solfège and theory, performance is also best taught as means to an end, or else students lose sight of the original target.

1.1.2.2. Solfège as an educational challenge

Literature arguing that solfège places unreasonable demands on the beginning musician has already been presented (see sections 1.1.1.2 and 1.1.1.3 for a discussion of solfège in England and the USA); many music educators have claimed solfège to be too difficult, too boring or too time-consuming for students, therefore of little use. Once more, the question arising is whether or not solfège, or any other music-reading or music-learning system, can be taught in isolation, as a supplementary tool within the framework of a very specific music education system, as a means to an end or as an end in itself; the central issue underlying all these questions is whether or not music education in a society should be part of a centralised, detailed programme, or a flexible entity, adjustable according to each community's characteristics and needs at a given time. In the latter case, when the exact content of a music lesson in a classroom is subject to the educational ideology, skills and personality of the individual music teacher, the main share of responsibility falls on the higher education institutions responsible for teachers' training; this approach does not guarantee homogeneity in the quality of the resulting music cultivation (Swanwick, 1988). On the other hand, a centralised music education policy is usually connected with the more 'traditional', or rigorous, approaches to music education (Swanwick, 1988): such rigorous methods, combined with the accessibility of all kinds of music by contemporary school students, often give rise to a 'quaint musical subculture' (Swanwick, 1999a, p. 127), which results in the undesirable effect of music being regarded as a peculiar and even elite activity. On the opposite side of the spectrum stands what Swanwick (1999a) determines as 'eventfulness' (Swanwick, 1999a, p. 133), a characteristic that contemporary music education should feature in order to make the music lesson and, subsequently, music itself appealing to students. It is interesting to note Swanwick's

(1988) critical stance regarding solfège and, more specifically, the Kodály method: he asserts that it constitutes an obsolete tool, inappropriate for the teaching of music outside a rigorous theoretical framework. As the purpose of the present thesis is to examine specialised characteristics of solfège as a note-naming system, rather than look at the potential of its incorporation in a particular music education system, it is outside of the scope of the present study to review the advantages and disadvantages of specialised solfège methods: from the ethnomusicologically-sensitive, moveable-do method of Zoltán Kodály (Dobszay, 1972), the innovative system of Carl Orff (Shamrock, 1997), who devised new instruments that fit best his pedagogical initiative, to the holistic, movement-based Eurythmics of Émile Jacques-Dalcroze (Seitz, 2005) and the keyboard-based approach of the Yamaha system (Miranda, 2000), a shared characteristic of all these methods is the importance they attribute in the acquisition of pitch names in very early stages of music learning. As the purpose of the present study is the empirical investigation of fixed pitch names in themselves, independently of the framework they have been taught in, no elaboration will be attempted on the comparative merits of the above systems; after the empirical grounding of solfège effects is achieved, its incorporation in a general music education philosophy will be the subject of further research.

Swanwick's criticism of the Kodály method, however, could be extended to a more general criticism of solfège; the characteristics of solfège, especially the use of unique syllables for note names, provide it with the potential to function both as a contributing factor to the creation of the 'quaint musical subculture' Swanwick describes, but also as the basis for a creative, 'eventful' approach to music. Characteristic accounts of such cases are given by Choksy (1969), who describes the rise and fall of the Kodály method in the USA as a result of its introduction by advocates of the 'rigorous' approach to music literacy; Shamrock (1997), on the other hand, gives an enthusiastic report of an amalgamation of Kodály solfège with the Orff-schulwerk method, while Rogers et al (2008) present the application of a singing-based pilot programme in elementary schools, loosely based on the Kodály method; Houlahan and Tacka (2008) also present the Kodály method from a contemporary, cognitive perspective. All the above studies emphasize the value of singing as a decisive factor in the success of the solfège-based method they are examining; solfège, unlike other music-reading systems, is inextricably linked to music performance through song. The value of singing, especially in the early

stages of music learning, has been repeatedly stressed by researchers (Swanwick, 1988; Stewart and Williamon, 2008), either in relation to the cultivation of general musicality and confidence in the musical life of individuals (Sloboda et al, 2005), or in relation to the parallel improvement of specific musical skills, such as music reading (Galin in Blasius, 1996, p. 10), aspects of instrumental performance (Morphy, 1918; Kendel, 1955; Hood, 1960; Drake et al, 1999; Wolbers, 2002) and general aural skills (Richter, 1938; Fine et al, 2006).

Another critique of 'solfège systems' in general – the word 'solfège' here denoting any use of arbitrarily chosen syllables as opposed to numbers or letters – illuminates a different aspect of the debate: if and how solfège can meet the demands of music education in transient societies (Brown, 2003). According to Brown, the problem with solfège is purely practical as it requires more classroom time than do numbers, letters, or a proposed 'mixed' system, which the author puts forward as the most appropriate solution to cover the needs of already cluttered curricula as well as the needs of children who move homes and schools and have to constantly adapt to different systems. A similar argument, though less elaborated, appears in an earlier article, whose author argues that instrumentalists are the only good sight-singers as they can rely on kinaesthetic memory to pitch notes correctly (Adler, 1975); solfège and similar systems are considered to be much less reliable due to their arbitrariness. Another alternative note-naming system, the 'tone-word' system has been proposed by Carl Eitz (reviewed by Jones, 1960): Eitz's system is based on the same principles as solfège but using different syllables, as well as entirely different syllables to denote sharpened and flattened notes. This system, however, is strikingly uneconomical as there is a total of twenty-one different, completely unrelated syllables to learn in order to be able to sing any given piece; it is not clear why solfège syllables, a system that has been successfully tested for centuries, should be replaced by different, yet equivalent syllables. The same argument is also made in Hughes (2000), who defends the already existing pitch-naming systems in a variety of cultures as being highly functional and governed by perceptual universals; a similar conclusion is reached in the empirical investigation of children's musical organisation abilitites by Brand (2000). Moreover, an early study by Gerson-Kiwi (1958) supports that even the more 'rigorous' solfège systems, like the Kodály, can be adapted to fit the needs of different societies and cultures.

The problem of accommodating different musical and cultural traditions might sound as quite unique to specific countries or areas with a particular social profile; it can be generalised, though, to be seen as a potential problem for the international music community: the use of a specialised system for referring to music, such as solfège, automatically excludes, at least to a certain degree, individuals who are not familiar with or fluent in this system. In order to verify this statement, the questions that need to be answered are the following:

- Is it true that solfège syllables are more arbitrary and therefore harder to teach than letters or numbers?
- Which are the characteristics that a music-referential system must have in order to be appropriate and functional as a universal code?

The above questions are both a matter of perception; the issues of if, how and to what extent solfège is linked to human perception of music are going to be discussed in the following section.

1.2. Solfège as perception

At the time of the completion of the present thesis, solfège and perception had never been looked at in combination by researchers, at least not directly. Many assumptions can be made about why solfège is neglected in cognitive research: the dominance in the literature of researchers from countries which do not include solfège in their curricula, the reluctance of researchers familiar with solfège to re-examine such a traditional or, for some, even obsolete method of music teaching, in fear that it could be interpreted as an attempt for the reinstatement of rigorous, old-fashioned classroom teaching, or simply the rapid advances in music cognition that make a potential comeback of solfège seem untimely could be some of the reasons for this neglect. Another plausible possibility, however, could be that solfège occupies an extremely ambiguous space in the universal mind, and in the minds of musicians it falters between being conservative or creative, alpha or omega, backbone or accessory, music or language. For this reason, the next sections will introduce the idea of solfège as a matter of perception first by

setting out the relevant theory and creating analogies with other domains, in order to illustrate that solfège and our perception of music are actually very closely interrelated. Literature linking, either explicitly or implicitly, solfège to perceptual processing will be presented in subsequent sections and, in the last section, the question of whether solfège is processed as music or as language is going to be theoretically tackled.

1.2.1 Empirical observations

It would be helpful to make some empirical observations based on anecdotal evidence in order to gain an initial understanding of solfège seen as a perceptual process. For this purpose, it is useful to divide the population of musicians into three broad categories: musicians actively practising solfège from the beginning of their music education; musicians who have never come across solfège; musicians who have come across solfège at some point in their lives but do not use it, usually due to lack of adequate experience with it. It is interesting to consider the reactions of the first group towards the notion that the other two groups even exist: most musicians familiar with solfège cannot even begin to think how it is possible to conceptualise music without knowing the names of the notes; letter names provide a somewhat equivalent point of reference, but still the notion of not being able to *sing* a pitch simultaneously with assigning it its verbal label always strikes solfège musicians as being extremely bizarre and even as being a kind of impairment. Such strong reactions resemble people trying to imagine how it would 'feel' and how they would be able to perceive the world if they were congenitally blind or if they lacked knowledge of names concerning a specific perceptual category. The most common examples to illustrate the necessity of labelling in the visual domain come from colour perception, as in the case of painters and art critics using dozens of different words for what the layman would simply call 'green', or the arguable existence of nine different words in the language of Inuit for 'white'. What happens to the non-visual artist, non-Inuit in the above instances, is that they are not able to distinguish effectively and/or verbalise such a distinction between hues of green or white because of lack of the appropriate linguistic labels for them in their vocabulary.

In line with the above observations, the empirical evidence indicating that people trained in solfège cannot cognise musical reality without using solfège syllables points to the empirical conclusion that solfège is, indeed, a matter of perception: people who practise it seem to have it ingrained in their musical functioning as a perceptual attribute of music. This type of empirical evidence provides an appropriate starting point for a perception theory of solfège: all theories of perception are born the moment someone gains conscious knowledge of perceiving reality in a particular way. On the other hand, empirical observation is not evidence enough, as one can argue that the case of solfège resembles that of optical illusions: it involves a piece of knowledge acquired in such an early developmental stage that it is almost impossible to abort; instead, all reality is subsequently modified accordingly to fit in the pre-constructed perceptual frame. Following the example of optical illusions, there are individuals that can experience reality – in this case, musical reality – without knowing solfège, and most of them are very likely to argue that solfège is nothing more than an alternative, dispensable pedagogical approach, which can be of no use to already accomplished musicians. Where does the truth lie between these two empirical approaches, and how can we discover it?

Speculation and theorising on the topic is most likely not going to bear any more fruit, as we are tackling a matter of perception; perception has to be experienced, as conscious knowledge cannot be simulated. Instead of theory, it is useful to observe the formation of perceptual, behavioural and emotional attitudes of musicians with no prior knowledge of solfège, who have received systematic and extensive training in it. Only they can be in a position to contemplate the nature of their new asset and its consequences and applications realistically, logically and, to an extent, objectively. This type of empirical research can establish a primary grounding that will ultimately lead to a complete understanding of solfège functionality within human cognition. In terms of empirical research in relation to solfège, Martin (1991) conducted an interesting experiment in a sample of 65 first-grade students: results showed that, after a year's instruction, the instruction method – solfège syllables, letter names, or numbers – did not have a significant effect on the students' performance on simple dictation and sight-singing tasks. This experiment, however, has not been replicated or followed up; moreover, it is not clear whether the teaching part of the experiment revolved exclusively around the experimental pentatonic materials or if it was a self-sufficient

music instruction programme according to the Kodály method; more detail about the procedure followed is necessary in order to evaluate thoroughly the outcome of this piece of research.

Before proceeding to experimental investigation, the first step is to provide a theoretical grounding for the perceptual relevance of solfège within a hybrid musical/linguistic referential system. In order to do this, the following questions need to be answered:

1. Which characteristics of solfège formulate its dual (musical and linguistic) nature?

2. What makes solfège perceptually and functionally appropriate in a linguistic referential system?

3. What makes solfège perceptually and functionally appropriate in a musical referential system?

4. How can solfège be accommodated in contemporary theories of perception and cognition?

1.2.2. Linguistic and musical traits of solfège

According to the definition provided in the introduction of the thesis, solfège is 'the practice of assigning the verbal labels 'do', 're', 'mi', 'fa', 'sol', 'la' and 'si/ti' to particular acoustic frequencies within a musical context'. This definition makes it clear that solfège is always used in relation to music as an art form, or music as defined, very broadly, in the New Grove Dictionary of Music (Nettl, "Music.", *Grove Music Online*); it was conceived in order to cover a gap in music theory at the beginning of the 11th century and, in this sense, is no more and no less a part of music than notation. Just like music notation, solfège does not purport to be intrinsically tied to music as a natural phenomenon; since, however, it has emerged from a Western society to be adopted, at least for a certain amount of time, by all members of this society in order to serve its music, it can be argued that it is intrinsically linked to Western music as a social, anthropological and cultural phenomenon.

On the other hand, there are two main points which prove that solfège is a part of language: the presence of solfège syllables in all general (i.e. not only music-specific) dictionaries and their functional compliance with the rules set by both dominant approaches to language, the Saussurean and the Chomskyan. According to the former, solfège would be defined as a system of signs utilised by a specific community, in which arbitrary linguistic labels are attached to specific objects or ideas (Saussure, 1983). According to Chomsky's Generative and Transformational Grammar (Chomsky, 1957, 1971) the notion that solfège would be part of a semiotic system remains the same; what changes is the processes through which the system is acquired by the speakers, as well as the potential classification of solfège in a syntactic category. Formal analysis of solfège according to either the semiotic theory or the generative and transformational grammar is beyond the scope of the present research, and, from a musical standpoint, rather irrelevant.

1.2.3. Linguistic relevance of solfège

The fact that solfège syllables are part of the human language provides enough proof for their linguistic appropriateness; there are, however, two notable points about solfège's linguistic identity:

1. The issue of isomorphism. Isomorphism in language, i.e. the one-to-one correspondence between words and specific objects (Παυλίδου, 1997), raises the question of which came first: the perception of the object as a separate entity or the ability to define it as such through a word, which then leads to its conceptualisation as a separate entity. Before they were assigned their solfège names, notes existed as physical objects but were not categorised; discriminations were made on the basis of relative similarities and differences. When Guido d' Arezzo introduced the innovations of staff notation and solfège syllables he eliminated much of the obscurity of neumatic notation and formulated Western music as we know it; it is plausible to assume that, for most people, the concept of the note became concrete after solfège, which provided the possibility of reference through the common linguistic functions of generalisation and abstraction. In this respect, moveable do was a natural first step towards the further refinement of solfège syllables to denote specific pitches instead of scale degrees. From this viewpoint, solfège is not only linguistically acceptable and relevant, but, being a

part of language, actually functions as an 'essential constitutive ingredient of human reality' (Barlett and Suber, 1987, p. 5) – in this case, musical reality.

2. Universality of solfège. The fact that solfège syllables were specifically created to serve as linguistic labels to musical pitch and that they are adaptable to fit every language's phonetic system render it a unique and rare example of true universality; this universality facilitates enormously the communication between musicians of different nationalities. Not everyone shares this view about the universality of solfège syllables: one of the oldest testimonies of distrust towards solfège comes from a book review published in the Musical Times in 1868; the author argues that correct pronunciation in singing can be best achieved through systematic practice with a language's own words, rather than using 'the questionable method of what is called "solfeggio" (Review of *Elementary Course*, 1868). This is refuted by papers which provide evidence on solfège being used as a vocal exercise for articulation (McNaught, 1892; Sands, 1943–1944) as well as from contemporary musical reality, in which vocal teachers around the world insist on the use of solfège. The perspective given in the 1868 review might at first seem naïve in its conception, coming from an author who is evidently slightly confused about the functions of solfège and moreover displaying signs of a somewhat irrelevant empathy towards the innate musicality of the English language; in closer examination, though, the argument in fact stems from a core issue concerning solfège: its arbitrariness and its musical appropriateness. The issue of practical, or phonetic, musical appropriateness will be addressed in the following section; first, the issue of arbitrariness of solfège syllables will be tackled.

That solfège is arbitrary, as opposed to letter names and numbers, is also argued in a much more recent paper by Brown (2003), who proposes an alternative syllabic system which he asserts is easier to learn. It is interesting that in Brown's paper solfège syllables are constantly being referred to as being unfamiliar and arbitrary, as opposed to the 'actual' letter names of the notes; in the same perspective, Brown argues that solfège requires a 'two-step' thinking process, of which step one is identifying the note and step two is assigning the correct solfège label to it (Brown, 2003). As Rappaport (2004) argues in his response to Brown, any system for labelling pitch can be perceived as being the 'actual', 'regular', 'original', or 'primary', as long as it is the one taught first. In terms of referential systems, there is no more arbitrariness in calling a musical tone 'do' than calling it 'C' or '1': all three function as symbols – in the Peircean sense

– and are assigned arbitrarily to the notes. One might argue that there is an implied gradation occurring in the latter two cases, in that we can conceptualise an ordered progression of letters, numbers and pitches as 'ascending' or descending'; the same sense of gradation, however, is no less true for a child that has been trained in solfège, who also learns to perceive the progression 'do-re-mi' as ascending and 'mi-re-do' as descending. In this sense, solfège syllables are not more difficult to teach or learn than letters or numbers; the only difference letters and numbers have compared to solfège is that, normally, a child is familiar with these two systems *in advance*. This, however, is not necessarily an advantage: on the contrary, the fact that solfège syllables only carry musical connotations renders them more intrinsically tied to music and therefore more appropriate than any other, pre-existing system.

1.2.4. Musical relevance of solfège

The appropriateness of solfège syllables within a music referential system was established in the previous section about isomorphism; Western music as we know it today started being organised for compositional and pedagogical purposes after, and partly as a consequence of, the introduction of solfège. On a more practical level, there are two main points, closely interrelated, which prove solfège's musical appropriateness, both in regard to the phonetic structure of solfège syllables.

1. Solfège syllables are extremely comfortable to sing, compared to both letter names and numbers; this has been proven throughout music history, as solfège has been used to improve the tone and enunciation of singers (Sands, 1943-1944). A common cadence in La, or A, minor/major provides a characteristic example: using letter names one would have to sing the rather uncomfortable sequence 'A-E-A', whereas the equivalent solfège syllables are 'la-mi-la'. An empirical survey that would investigate the relationship between the use of solfège and pronunciation accuracy in singing in various languages would be the optimal solution in order to clarify the way solfège syllables are adopted and being made functional by singers in various countries; the main arguments, though, opposing the belief that solfège is detrimental to the intelligibility of singing are the following:

- Solfège has been successfully used for centuries in a very large number of countries around the world (Chapter 1.1): in all these instances, the pronunciation of solfège syllables has been assimilated to each language's own phonetic system. For example, the syllable 're' is pronounced [ɹeɪ] in English, [Re] in French and [re] in Greek (phonetic transcriptions according to the International Phonetic Alphabet, International Phonetic Association, 1999).

- Many vocal tutors support solfège as being optimal for achieving clear articulation: the oldest testimony for such an argument comes from Domenico Corri, an 18th-century Italian singer who prospered in Great Britain as a conductor, singer and singing teacher (Sands, 1943–1944).

2. The singability of solfège syllables and, subsequently, the fact that they are almost invariably used in combination with singing make them extremely useful in various occasions when instant communication between musicians is required. Such is the example of a conductor who needs to verify that what is written on his score is the same as the viola part: using letter names he would have to 'describe' the passage by saying 'C quaver – staccato, quaver rest, C dotted quaver tied to A semiquaver, tied to F dotted crotchet – tenuto...' or sing the passage using a neutral syllable, stating the note he is starting from and relying on his fellow musicians' aural skills; things are made remarkably simpler and faster when singing the phrase using solfège.

1.2.5. Solfège in the brain

What has been stated above is partly echoed in Brown's paper (see discussion of the universality of solfège in section 1.2.3), who, in investigating the development of musical literacy in transient societies (Brown, 2003), argues that, contrary to instrumentalists – who make extensive use of kinaesthetic perception – singers learning music by rote do not necessarily have an internalised conceptual framework for pitch relationships. What would help singers conceptualise music would be the use of a singing system, such as solfège, or any other letter, number or other label in addition to the traditional notation system (Brown, 2003, p. 47).

Seen as such a system, solfège elements comply with the constraints for the formation of perceptual categories put forward by Roth (1986); according to these rules, cultural factors combine with innate attributes of objects and constraints of our perceptual system in order to produce meaningful categories, useable by all members of the group that operates them (Roth, 1986, p. 22). The innateness of solfège's musicality has just been argued in the sections above, whilst the issues of perceptual constraints as meant by Roth do not apply to solfège; the cultural factor, which resulted in the production of the meaningful perceptual category that solfège is, was the lack of an appropriate referential system for pitches in Western music at a certain period of time.

The concept of modularity as proposed by the philosopher and cognitive neuroscientist Jerry Fodor (1983) has incited researchers to investigate the possibility of music perception being completely distinct from other cognitive functions. According to the modularity theory, the brain can be physically and structurally separated in independent subsystems – modules – with unique processing, storage and functional features. Peretz has tested this hypothesis in a series of neural imaging studies (Peretz and Morais, 1989; Bonnel et al, 2001; Hébert et al, 2003; Peretz and Hyde, 2003; Peretz and Coltheart, 2003; Peretz et al, 2004; Lidji et al, 2007) to conclude that there is strong evidence that speech and music are processed by distinct neural substrates. This body of research also suggests that not only is processing separate for music and speech, but also separate for different components of music, such as tonal structure, temporal structure and pitch interval information. If this is the case and given that solfège occupies an ambiguous ground between being music and/or speech, it is certainly worth investigating the neuroscience of solfège perception and processing; such research would provide insights into both solfège cognition and, most importantly, music perception and cognition in general. Moreover, since modules are assumed to be informationally encapsulated so that there is no exchange of information between them, solfège would have to be primarily a function of either the language processing or the music processing module. If the hypothesis that solfège is a part of music processing is correct, it will provide an extremely strong piece of evidence in favour of the existence of a distinct, music-specific system, which utilises a physically separable set of neurons extending from the tonotopically organised cochlea to specified areas of the brain (see Brandler and Rammsayer, 2003, for a review of studies on the neuroanatomy of music perception; also Marin and Perry, 1999).

Although this type of research is yet to be conducted in direct relation to solfège, subsidiary findings that could be used to support the existence of solfège as a unique musical function, which employs separate processing subsystems, as well as evidence for the contrary, can be found in some of the studies examining singing and/or cases of congenital or acquired deficiencies affecting either language or music. Either in the cases of aphasia without amusia, in which the patient cannot recognise, recall or produce language correctly, while the music faculty has remained intact, or in the cases of congenital amusia with no language impairment, it has been shown that the combination of words with melody does not improve performance in verbal tasks (Hebert et al 2003; Peretz et al 2004). Although these studies are based on either single reported cases or extremely limited samples, Peretz et al (2004) argue that the study of single abnormal cases can definitely lead to valid generalisations on the normal population, as brain damage results in deficits which break down an extremely complex processing system, thus making it easier to investigate. Evidence from this body of research points towards an autonomous processing of words and music regardless of the seeming integration between the two in the case of singing, which has also been proposed by other researchers in cognitive neuroscience and neuropsychology (Samson & Zatorre, 1991; Perry 1993; Bonnel et al, 2001); if solfège is a direct equivalent of song, thus a purely linguistic element which is simply combined with melody in the last stages of production, then it will be expected to present the same effects of selective impairment when running a parallel linguistic task.

On the other hand, Peretz has also reported a case of a 'tone-deaf' individual who, despite his complete inability to reproduce musical pitches, either separately or in a sequence, even remotely accurately in other contexts, was able to reproduce sequences perfectly using solfège syllables (Peretz, 2007). The person had no perceptual problems in dealing with music and was classified as an absolute pitch possessor; his deficit was limited to production. The result was that, when asked to sing a familiar tune ('Happy Birthday') using the lyrics or a neutral syllable, his performance did not match the original even in its melodic contour; when sung with solfège syllables, however, performance was impeccable. So, although the man was classed by his social environment as 'tone-deaf', he was actually an absolute pitch possessor who, presumably due to lack of training and/or bad relative pitch skills, was subsequently unable to function musically in a social context. However, as this individual case was

not specifically investigated in respect to solfège – the person happened to be one of Peretz's research collaborators – it is possible that the selective preservation of solfège skills occurred because of individual characteristics such as the lack of musical training and possession of absolute pitch, rather than due to specific properties of solfège. A more detailed review of absolute pitch in relation to solfège will be presented in a separate section (Chapter 1.2.7) because of the unique nature of the skill and the rarity of its occurrence, which, in some cases, makes generalisations impossible.

The issue of the lateralisation of music is raised in all the above studies; as stated before, there have been numerous suggestions about possible areas in the brain where music would be – primarily – processed and stored. The general conclusion from brain research is that music perception is certainly less lateralised than language perception (Marin 1989, 1999); quite interesting in this respect are the findings of Perry (1993), who studied pitch memory in experienced pianists. He also concluded that there is the possibility of a separate neural substrate for pitch information; moreover, he found that melody processors in these specific circumstances were laterilised in the right hemisphere, rather than the predominantly language-related left.

Finally, several studies have shown differences in music processing between musicians and non-musicians (Bever and Chiarello, 1974; Brandler and Rammsayer, 2003; Fujioka et al, 2005). Earlier studies assumed that this difference reflected a particularity in the inner quality of music; later studies showed that the differentiation is a result of brain organisation as formed by early training. In respect to solfège, one might add that the fact alone that musicians learn to refer to the smaller units of music in a particular way can possibly result in a different coding, which will naturally generate a differentiated set of responses and a different mapping in the brain.

1.2.6. Solfège as singing

At this point, a clarification should be given regarding the link between solfège and singing and the extent of similarities and differences between the two. One could argue that solfège is not singing, since it most certainly falls into the category of non-propositional, non-generative speech, whereas singing with lyrics employs generative

language, albeit poetic. On the other hand, singing is very often classed as non-propositional language in experimental contexts on the grounds that the lyrics of familiar songs are extremely well rehearsed and thus produced automatically; a characteristic example is the *da capo* reproduction of a song in order to recall a particular verse or word in it, which reveals the existence of an 'echoic' rather than a 'pragmatic' memory trace.

What is the connection between solfège and singing, then? First of all, there is the obvious similarity, which is also the main reason of confusion between the two: they both use the vocal tract as means of production and they both result in a musical output. The difference lies in the cognitive processes resulting in this output: singing is based on a well-learned set of linguistic rules which have primarily other functions and whose product, the lyrics, is combined with music for expressive reasons; solfège syllables, on the other hand, have been learnt in order to serve music learning and are primarily used in combination with it – the referential advantages of using solfège, mentioned in section 1.2.4, are only a by-product of this principal function. In this respect, solfège syllables bear more resemblance to the words used for numbers: the words 'one', 'two', 'three' are undoubtedly a part of language, they also have referential and abstractive capacities, but our perception of them is always mediated by their mathematical values; the words are the vehicles which help us conceptualise the inherently non-verbal.

1.2.7. Directions from absolute pitch research

Absolute pitch (hereon AP) research provides extremely useful insights into the relationship of category labels and musical sound; on the other hand, it is not certain that these insights are not only relevant to the 0.01% of the general population that happens to possess AP (Takeuchi and Hulse, 1993; Ward, 1999). Because of this controversial aspect of AP, relevant research findings will be presented in this separate section and the implications these findings could have on solfège research and theory will be briefly discussed.

Researchers argue that the association between auditory stimulus and 'category labels' is necessary for the emergence of AP (Levitin and Zatorre, 2003; Levitin and Rogers,

2005, for a review); it is characteristic that in order to illustrate this association they often use the parallel between colour categories and pitch categories (see section 1.2.1). In addition, cognitive psychologists report that neural regions recruited in pitch recognition by AP possessors are the same as in conditional associative learning, which is general for verbal labelling of sensory percepts (Zatorre et al, 1998, in Levitin and Zatorre, 2003, p. 106); a possible verbal encoding or other cognitive strategy, arising from the importance of labelling pitches, is also argued in Chin (2003). This leads to the assumption that in the case of AP specific neural networks are tuned to pitch labelling; this is in line with research findings on the modularity of music perception presented above, as well as with the notion that solfège, seen as a category label for musical pitch, is music-specific both functionally and anatomically.

The importance of category labels is reflected in Levitin's two-component model for AP, which includes a pitch memory and pitch labelling element (Levitin, 1994; Levitin and Rogers, 2005); this is in accordance with research attributing the superiority of AP possessors to the verbal rather than pitch memory (Takeuchi and Hulse, 1993). Takeuchi and Hulse (1993), in reviewing literature on different measurements of AP acuity, report one measure to be the level of decay in memory for pitch. They report that, in music recall tasks, AP possessors' performance does not deteriorate after 1 minute, whereas at the same time non AP performance deteriorates to chance; this is interpreted as evidence that non AP possessors are solely based on echoic memory, whereas AP possessors have verbal memory readily available (Takeuchi and Hulse, 1993). The assumption about the utility of verbal labels in AP is demonstrated further in experiments where pitch names cannot be used in order to discriminate between pitches, as happens in cases above certain frequencies (over 5000 Herz) or in frequencies falling between tones; in such cases AP and non AP possessors perform the same (Takeuchi and Hulse, 1993; Ward, 1999). It would be worthwhile to test whether this kind of memory for pitch labels displays the same characteristics as memory for speech and phonological elements; this would clarify the assumptions regarding the extent of verbality of AP memory.

If AP is mainly reliant on verbal coding structures, then naturally the pitch labelling networks involved will be expected to comply with the same rules as other neurodevelopmental events (Levitin and Zatorre, 2003). This raises the issue of the

importance of early training in the acquisition of AP; researchers who advocate a primary role for training during a critical period (see Levitin and Zatorre, 2003 for a review) stress the importance not only of the *timing* in which musical training takes place, but also of the *type* of training. Researchers argue that musical training, in so far as it emphasizes pitch relations, reinforces relative pitch and can prove detrimental to the acquisition of absolute pitch (Takeuchi and Hulse, 1993; Ward, 1999; Levitin and Zatorre, 2003). Unfortunately, data from research relating solfège to the acquisition of either relative or absolute pitch are quite inconsistent and mostly based on theoretical accounts, observation and anecdotal evidence than on experimental findings; there seems to be a general agreement on the fact that solfège in most cases improves relative pitch (Smith, 1934; Bullis, 1936), but there is also evidence on solfège facilitating the acquisition of AP, with the most prominent example being that of the Yamaha method for music teaching, which explicitly uses fixed do solfège in order to cultivate AP (Miranda, 2000). This provides supplementary evidence that music training aiming to the acquisition of AP has to be specifically appropriate; emphasis on category labels, such as solfège, is the first step towards this goal, but the establishment of pitch names has to precede understanding of pitch relations. This is also in tune with the natural developmental sequence of sensory perception in general (Takeuchi and Hulse, 1993): children first learn how to perceive things absolutely and only later, with the introduction of abstractive thinking, are they able to form relations between perceptual events.

1.2.8. Summary

This section presented solfège as an element of human perception of music starting with anecdotal empirical observations on the perception of music with and without solfège, followed by a theoretical account of solfège's perceptual traits. In the empirical domain, valuable insights are gained through parallels with colour perception; theoretical analysis of solfège consists of an account of the factors which formulate the double – linguistic and musical – nature of solfège and justify its referential appropriateness within the two systems.

Evidence from cognitive neuroscience and AP research was presented supporting the existence of separate substrates involved in the perception and processing of music and language and the importance of verbal coding for the categorical perception of music; further research, monitoring the interaction between linguistic, musical and solfège tasks, is necessary in order to identify solfège as belonging to either of the two systems. In conclusion, the notion that solfège is a matter of perception, thus forming the basis of an overall different approach to music, seems irrefutable; the perceptual relevance of solfège provides a plausible explanation for its dominance, for many centuries, in music education curricula internationally (section 1.1). It is exactly the dual nature of solfège that renders it optimal for a thorough understanding of fundamental musical structures through a familiar –linguistic – medium; for all the above reasons, solfège deserves more attention both from the part of music educators and students and from the part of researchers, who can help by illuminating the functions of solfège within the framework of human cognition. This, of course, does not entail that solfège can be beneficial when practised in isolation; like all other elements in music learning, it can function most effectively and yield excellent results only when combined with a general approach to music which is time, place, material and subject appropriate.

1.3. Solfège, memory and memorisation

It is important to settle from the beginning the difference between music (or musical) memory and music (or musical) memorisation and define the relationships between them. The most important feature of music memorisation, which differentiates it from the general notion of musical memory, is that it is a deliberate action, taken usually by trained individuals towards a specific goal. Musical memory, on the other hand, is memory *for* music, which, in the relevant field of research predominantly refers to melody recognition; researchers have not yet come to an agreement whether this involves a different process for musicians and non-musicians and whether it is controlled by music-specific memory stores or not. Undoubtedly, there are issues common to both music memorisation and music memory: general and musical memory capacity, storage and processing functions naturally affect the processes of music memorisation, the techniques musicians use when memorising music, the order and manner in which they employ these techniques, and the time they take in order to complete a memorisation task.

In the case of solfège, which functions as a verbal expression of musical data, relationships between verbal and musical memory and verbal and musical memorisation are even more intertwined and complex; an empirical investigation of these issues will be presented later on, in Chapters 3 and 4. What has to be clarified, first of all, is the proposed relationship of solfège with human memory; after the connection between solfège and music memory has been established relevant literature will be reviewed.

The assumption posited in the present thesis is that **solfège has a double function, owing to its combined verbal and non-verbal nature**. The first function is active during the memorisation process, and the second function involves retrieval processes.

First function - Memorisation process: The first function bears a direct relationship with the perception of the musical stimulus (see section 1.2 for a discussion of solfège in relation to perception). Whilst in the act of memorising, musicians have to encode the material in one or more ways and store it so that it is easy to retrieve on demand. This study proposes that solfège provides a more 'solid' ground for a memory trace, as knowledge of solfège requires the existence in long-term memory of a separate **non-**

sensory, **symbolic** trace for each single note. If this is the case, the encoding process is executed on different terms than when there is no established solfège knowledge as there is an additional, extremely stable and well-rehearsed point of reference. This can be better illustrated by using, once more, an example from the visual domain, drawing an analogy between a musician trying to memorise notes in a short phrase and a painter trying to memorise the order of colours in a picture of a rainbow. The equivalent of trying to remember the musical phrase without solfège would be the painter not having the names of the colours readily available in his/her mind; in this case s/he would try to remember their order by preserving the sensory, visual trace as well as combining it with previously acquired information about the physics of light, the wavelength of each colour in the spectrum and so on. Being able to assign names to the colours does not necessarily make the memorising task easier or performance in it more accurate; it does, however, provide an unquestionably strong aid. The combination of sensory input, world knowledge and a symbolic-linguistic connection to the image provide a more solid trace for the rainbow; the same is proposed to be taking place in the case of solfège.

Second function - Retrieval process: If a process similar to the one described above takes place in the case of solfège and music memorisation, retrieval mechanisms for music are subsequently modified as well; this conclusion is based on research showing that memory trace formation depends on perceptual processing having taken place previously, as well as on the spread, depth and modality of encoding (Craik & Lockhart, 1972; Craik and Tulving 1975; Eysenck, 1977). In addition, since the depth of encoding is also crucial, it is plausible to assume that solfège provides not only an additional – verbal – manner of encoding, but also an encoding relying on deeper structures: as pitch is imbued with endogenous semantic meaning, the sensory input is being 'promoted' to the category of a signified.

In spite of the clear-cut division between memory and memorisation, however, it is research on the former that provides the necessary links between cognition and memory processes, which, in turn, feed into memorisation techniques. Thus, implicit evidence that solfège actually functions in the twofold way proposed above can be found throughout the whole body of literature on memory theories; this seemingly incoherent lot of information will be presented in two sections, regarding general memory and

working memory theories respectively; the last section will review research on music memorisation.

1.3.1. Memory theories and approaches

A very common debate in memory research is whether memory is a structure or a process (Hirst and Pinner, 1996, p. 244). Eysenck argues that this is primarily a question about the nature of information processing: the processing is mediated by *structures*, which are modality specific and provide a temporal ordering to the flow of information; these structures have as their input both processes starting with perception, like coding, as well as processes involved in the formation of strategies of recall (Eysenck, 1977). Numerous models have been proposed regarding the possible structural elements of human memory: long and short-term memory (hereon LTM and STM, respectively), working memory, episodic and semantic memory, and many more; what is important is that the dominant contemporary models are all a combination of multistore or modal models (Murdock 1967), based on the concept of different stores bearing different retention characteristics, and the notion of *levels of processing* first proposed by Craik and Lockhart (1972). Levels of processing, although not a theory per se, provided a conceptual framework for memory research thereon; it is based on the idea that the memory trace is a by-product of perceptual analysis. Thus, the different stores in the modal models reflect different levels of processing of the stimulus; the associations each stimulus generates determine the spread and depth of its encoding and, subsequently, its potential for retrievability (Craik and Lockhart, 1972; Craik and Tulving, 1975). In this perspective and bearing in mind the evidence presented in the previous chapter about solfège being primarily a matter of perception, it can be assumed that solfège offers the opportunity to process a melody in an additional way, adding to the spread of encoding; moreover, the new encoding is unique in that it is the only chance to link the physical aspect of the melody – pitch – with a specifically pre-determined verbal label.

This advantage solfège can offer is also supported by the notion of *encoding flexibility*, which refers to the ability to encode a particular stimulus in several different ways (Craik and Lockhart, 1972; Eysenck, 1977); encoding flexibility ensures that the

additional encoded structure generated by solfège will be useable and moreover helpful in order to recall a melody. The same point is carried further by Baddeley (1999) who argues that memory performance improves when the act of memorisation is combined with an activity: based on the ability for multiple encoding, this provides an additional argument in favour of the use of hand signs during the process of learning solfège syllables.

On the other hand, the *encoding specificity* (Tulving and Thompson, 1973) and *remembering operations* (Kolers, 1973) principles support the argument that encoding and retrieval processes may be similar or even identical; this means that encoding processes are carried out primarily to serve perception (Craik and Lockhart, 1972), and that 'retrieval processes represent the cognitive system's attempt to reinstate the same pattern of cognitive activity again at the time of recollection.' (Craik, 1983, cited in Craik et al, 1998, p.61). Further support for this comes from the field of cognitive neuroscience where the same neural networks are recruited for both the initial perception of information and for its retrieval (Squire et al 1985; Mishkin and Appenzeller, 1987; Moschovich et al, 1995). According to the above, if solfège is involved in the perception and encoding of music, then successful recall will need to involve recall of solfège syllables; the advantage solfège syllables bear over letter names as well as over kinaesthetic or structural information has already been discussed extensively in the present and in the previous section.

At this point, it is important to refer at least briefly to alternative models of interaction between verbal labels and retrieval, such as the *representational shift account* (Lupyan, 2008b). Lupyan shows strong evidence that category labels may not affect the initial encoding of an object, but rather distort its stored representation, with subsequent effects on recall performance patterns. Although this is a very important finding, with numerous practical implications on memorisation and recall theories, its impact on solfège perception theory is minimal, mostly due to the fact that such a theory is still in an embryonic stage. Theoretically, it is plausible to assign solfège syllables to either the group of specified object names – which is the main function of fixed do as opposed to moveable do syllables, as argued previously – or to the group of generic category labels – as happens in the case of both moveable do, but also in fixed do, if the fact that fixed do syllables still encompass a considerable amount of variations in octave range, timbre,

and so on, is taken into account. Experimental investigation to determine the exact nature and functions of solfège as a cognitive mechanism is necessary in order to enhance further a solfège perception theory and provide a detailed account of its links with music, pitch and language perception.

Moving back to general memory theories, another common distinction is made between episodic and semantic memory, which reflects the difference between memory for events and memory for the products of organised knowledge and encoding processes; in the first instance the stimulus can be raw sensory input, whereas semantic memory involves the retention of cognitive referents, usually mediated verbally (Tulving, 1972; Eysenck, 1977). Once more, solfège can potentially have an immediate effect on the formation of a memory trace for a particular piece of music: without the verbal encoding solfège provides, the music itself can only be part of episodic memory, and any semantic memory referents can only regard extra-musical or meta-musical elements, such as the structure of the piece, its harmonic analysis, or any interpretive or ad hoc semantic attributes. By using solfège, the substance of the music itself acquires a semantic referent which can be recalled independently of any other peripheral information.

A determining factor in the quality of retention is the amount and type of rehearsal taking place between perception and recall, especially in the period before the datum goes to long-term memory storage. As expected, retention varies directly with the amount of rehearsal and repetition (Craik and Watkins, 1973); rehearsal modes can be further divided into categories: simple maintenance, elaborative rehearsal and repetition (Craik and Lockhart 1972; Craik and Watkins, 1973; Eysenck, 1977; Segalowitz et al, 2001). In the act of music memorisation, solfège can play a double role: it gives the opportunity for additional rehearsal, which can be autonomous from the instrument or the lyrics, therefore free from any kinaesthetic or other connotations; furthermore, if solfège is practised regularly enough, the possibility arises that certain common motifs, such as arpeggios, scales or piece-specific motifs, will be crystallised acquiring a separate quality as a group, thus providing the potential for instant recall. The importance of the latter function lies, once more, in the fact that the encoding of the motifs will be encoding of their pitch information rather than of any other attribute; however, as this applies predominantly to tonal music, and only in cases when solfège is

in the core of one's music education from a very early stage, emphasis on this particular advantage could result in false generalisations.

Finally, memory for music is often discussed in connection to the 'skilled memory theory' (Chase and Ericsson, 1982). According to this theory, memory skill is complex and acquired through extensive practice; appropriate practice is assumed to lead to faster storage in LTM rather than extend STM capacity (see Aiello and Williamon, 2002, for a review of skilled memory connected to exceptional memory for music). Moreover, Krampe and Ericsson (1995) claim exceptional musical memory to be, along with absolute pitch, one of the defining factors of musical talent; possible connections between absolute pitch and musical memory have been discussed previously, in section 1.2.7. One aspect common to individuals displaying exceptional memory is the use of verbal coding regardless of the nature of the stimulus to be memorised; verbal, syllabic and phonemic sequences and matrices are constructed as part of a mnemonic strategy, in order to group information in a format which is presumably more convenient for recall (Ericsson, 1985). Mnemonics were originally regarded not only as a useful skill but also as a way of understanding the world (Hutton, 1987); in the same manner, solfège syllables constitute an easily learned phonetic code which can be used, as is its primary function, to perceive music altogether. The properties of good mnemonics are: 'ease of acquisition, continued usage, fit of mnemonic to task, flexibility and ease in combining mnemonics, and effectiveness of mnemonics in promoting long-term retention.' (Wilding and Valentine, 1996, p. 400). The issue of ease of acquisition in the case of solfège can be addressed in two separate instances: acquisition at an early age and later acquisition; although there is abundant literature arguing that solfège is most convenient and effective a way for a child to learn music (see section 1.1), no research has been yet conducted on adults learning solfège. The age when solfège is learned is also directly related to the issue of continued usage; the fit of solfège to music has been argued extensively on chapter 1.2 and the effectiveness of solfège into promoting long-term retention is yet to be proven, although there is strong theoretical evidence to support this. The fact that there is anecdotal evidence on musicians using solfège as a mnemonic code makes it plausible that it could be also regarded as a mnemonic strategy; however, stressing solfège's properties as a mnemonic strategy does not do justice to what it can really offer to the learning musician. The fact that it is formed optimally in order to serve its primary purpose should be best regarded as a means to an end and an

additional advantage rather than its primary function and aim; solfège is a complete system of learning, perceiving, understanding, reproducing and generating music.

1.3.2. The Working Memory Model

The working memory model, first proposed by Baddeley and Hitch (1974), has been one of the most influential and widely accepted models for explaining interactions between various memory stores, perceptual processes and attention. It is based on the notion that the short-term store is not simply a temporary store with limited capacity, but is itself constituted by separate subsystems with both processing and storage functions, essential to activities such as comprehension and reasoning in language, and sight-reading in music (Aiello and Williamon, 2002). This short-term processing and storage system is not independent from long-term stores: in order to overcome the restrictions posed by the relatively small capacity of short-term stores, individuals have to resort to using previously learned rules and strategies, stored in LTM, in order to chunk and categorise information (Ericsson, 1985; Ericsson and Kintsch, 1995; Berz, 1995). The constituents of working memory were initially proposed to be three: a central executive component with two slave systems, the phonological loop and the visuo-spatial sketchpad. The phonological loop is responsible for the manipulation of speech and other acoustic information and comprises a phonological store and an articulatory loop, responsible for controlling subvocal rehearsal; the visuo-spatial sketchpad deals mainly with visual data. The general supervision, manipulation and regulation of the flow of information between the two slave systems and towards the long-term memory store are performed by the central executive system. In 2000, Baddeley revised this model to include a fourth element: the episodic buffer (Baddeley, 2000; Repovs and Baddeley, 2006); this buffer functions along the phonological loop and the visuo-spatial sketchpad as a multimodal store for combined information from the other subsystems and from LTM. Since its emergence, the working memory model has been dominant in the field of human memory research; numerous extensions, clarifications and amendments have been proposed on its basis in order to accommodate phenomena which, although justified by the presence of working memory, would require the existence of more complex interrelations between coding, rehearsal, retrieval processes, attention and long-term stores in order to be fully explained. Arguably the

most common criticism of the working memory model has been its inadequacy in explaining the handling of acoustic information other than speech, as well as other sensory information, such as smell and taste; the combinatory nature of the recently added episodic buffer is not always successful in explaining certain specialised cases of memory feats and pathologies.

How and to what extent, then, does the working memory model apply to music? Berz (1995) provides a review on literature showing that musicians draw on LTM strategies developed through training for recall; the recall process, Berz assumes, must involve a specified musical processing component, connected to the memory task. Subsequently, Berz reviews research investigating working memory in music, trying to define whether working memory for music has the same characteristics as working memory for other acoustic information. Research findings show that, although there are phenomena in short-term memory for music which display similar characteristics to verbal STM, such as articulatory suppression, other common grounds in auditory STM, such as modality, primacy and recency effects, have been found to differ between music and language (Berz, 1995). The phonological loop in the working memory model is, by definition, restricted to verbal coding of acoustic information; since it cannot account sufficiently for music storage and processing, Berz puts forward the proposition to incorporate an additional slave system to the central executive, specific to music. The existence of a separate neural substrate for pitch information which supports working memory is also supported by Perry (1993) in a test with experienced pianists. An important indicator of the existence of such a specialised subcomponent is the unattended music effect (for a review see Berz, 1995). Unfortunately, there is no literature on unattended music running a parallel music task (Berz, 1995); on the other hand, research findings regarding performance on reading comprehension, verbal or phonological tasks while listening to background music are contradictory: performance is not influenced or is improved with instrumental music, whereas it is impaired with music with lyrics (Henderson et al, 1945; Martin et al 1988; Salame and Baddeley, 1989). This indicates that music combined with lyrics is making use of the same processing system as verbal tasks (Berz, 1995); the nature and extent of the interferences, however, render the three-component working memory model inadequate for the explanation of these effects. Even with the incorporation of the episodic buffer in the system, music processing seems to place extremely specific demands in storage and retrieval to be accommodated

by a general, multimodal, multifunctional and limited capacity store; Salame and Baddeley (1989) have also acknowledged the possibility of an additional form of acoustic storage capable of dealing with non-speech stimuli (Salame and Baddeley, 1989, cited in Berz, 1995, p.361). In the case of solfège, the extent to which it is similar to music with lyrics remains to be shown, as solfège is music-specific so it lacks the exogenous semantic connotations which would impair performance in linguistic tasks; experimental testing on solfège task performance with unattended normal and nonsense speech is necessary in order to shed light on these interactions.

Berz leaves the question regarding the nature of the relationship between the proposed music-specific subcomponent and the phonological loop open-ended; he states that the music subsystem could either be a specialised function of the phonological loop or an entirely different entity. In any case, he argues that the ties between the two systems ought to be extremely loose, given the differences between verbal and musical stimuli and the research findings showing differences in the processing, retrieval and decay of these stimuli. Berz also assumes that the nature of the musical component should be very similar to that of the phonological loop, also displaying a musical store, parallel to the phonological store, and a set of control processes based on inner – musical – speech (Berz, 1995).

The proposition put forward in this thesis is that there is, indeed, a distinct, music-specific system, which, however, is definitely not a part of the phonological loop; it is separate, and, just like the phonological loop, its formulation is dependent on developmental and evolutionary processes, environmental stimulation and appropriate training. The existence of the phonological loop is indisputable and presumably a result of the continuous presence and predominance of language as a perception, coding, understanding and communication medium, both in developmental terms and in respect to human evolution history. Music also seems to be inherent to humans, appearing invariably in all civilisations and also apparent in every human as potential (Merriam, 1964; Blacking, 1976; Gardner, 1983; Trehub, 2003; Mithen, 2005; Edwards and Hodges, 2007); unlike language, however, it has had numerous primary functions, expressive as well as performative, whose boundaries have not always been defined with precision. The unambiguous effects of music interference with other cognitive tasks, mentioned above, are apparent due to its continuous presence in human history; the functions, however, taking place in the cases of trained musicians are much more

specific and differ depending on the individual and on the training. In this sense, the same can be true for all other perceptual modalities; we can assume that, depending mainly on extensive practise, one could develop an exceptional olfactory working memory, or a specialised memory for taste. This kind of expanded modal organisation of working memory, which includes specialised units for modalities other the ones designated in working memory models proposed by Baddeley et al (Baddeley and Hitch,1974; Baddeley, 2000; Repovs and Baddeley, 2006), is supported by recent psychological studies (Andrade and Donaldson, 2007; Johnson and Miles, 2009). Andrade and Donaldson (2007) use short-term memory tasks to test odour memory in relation to odour recognition and digit recall, to support the possible existence in working memory of a subcomponent dedicated to the manipulation of olfactory information. Johnson and Miles (2009) examined comparatively order memory for odours, unfamiliar faces and pure tones; they found that participants were able to make absolute memory judgements for odours and that odour memory was the only one not affected by serial positioning, which also point to the conclusion that working memory is a modally organised system, with specialised components for certain modalities, one of which could possibly be olfaction.

Working memory can be thus viewed as a loosely defined set of functions, which are controlled by a central executive; the potential for development of a language, spatio-temporal, visuo-spatial, or music-specific subsystem is inherent as a kind of memory 'competence' in the Chomskyan sense: the possibility for such functions is present in everyone and the realisation depends largely on individual and environmental differences. If this is the case, a music-specific subsystem would be formed according to available perceptual and coding resources; its relation to the phonological loop would not be one of attachment but rather of sharing from a common pool of processes, as shown in Figure 1.1:
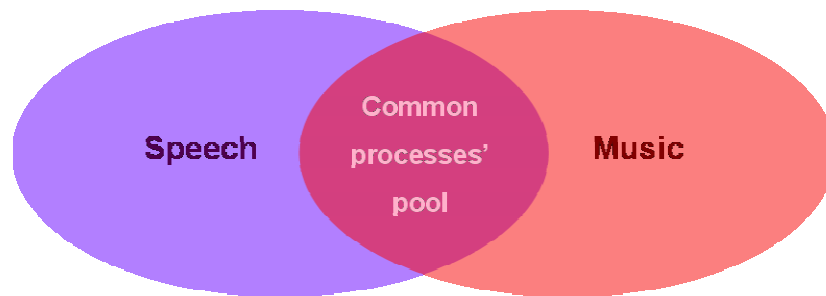
Figure 1.1: The purple overlapping area demonstrates the cognitive processes shared by verbal and musical functions.

The above are partly echoed in the definition of working memory capacity as a system of temporarily activated long-term stores combined with controlled attention (Engle, 2001); this proposition is based on the same percept of working memory being a flexible system, depending on or sharing resources with other cognitive structures. A similar point is made by Ericsson and Kintsch (1995), who argue about the existence of a 'long-term working memory', formulated by the skilled use of long-term memory storage (Ericsson and Kintsch, 1995, p. 211). This view, combined with the skilled memory theory (Ericsson, 1985) implies the possibility for the development of a domain-specific set of rapid retrieval skills, which would correspond to the working memory for this specific domain. In the case of music, if solfège is included in early training, we can assume that it promotes the use of a set of coding and retrieval processes common to the phonological loop for verbal information; rapid retrieval can be made possible through the use of all available encoding resources, through the close tailoring of the manipulation of these resources to serve music-related activities, as well as through use of familiar structures common to the verbal and musical phonological loop, as suggested above.

Although there is no direct evidence that working memory for music functions as a distinct subcomponent which can share resources with other subcomponents, namely the phonological loop, there is plenty of research on verbal aspects of human memory,

whose findings can be used to make some inferences about musical working memory. Neuroimaging research on the phonological loop and the visuo-spatial sketchpad confirms that these subcomponents of working memory are relatively independent; this makes it plausible that the same should be true about other modalities of perception. The fact that verbal material, and subsequently verbal memory, are predominant in research, can be attributed to the importance of verbal coding as well as to the convenience verbal material presents in its manipulation in experimental contexts (Baddeley, 1999, p. 7); at the same time, this is an argument favouring the importance of verbal coding of music through solfège syllables on the one hand, while it also reveals the potential for future research: the verbal nature of solfège renders it suitable for investigation using the same, accredited methods as the ones used in verbal memory research. The importance of verbal codes in music is also mentioned by Berz (1995), who argues that, if verbal labels are readily available, they create a relative advantage in performance in certain tasks, such as music dictation.

1.3.3. Music memorisation

The links between music memorisation, working memory and solfège may still seem obscure at this point; in order to clarify these relationships, the relation between working memory and music memorisation will first be explained and then different categories and techniques of music memorisation will be presented.

Working memory is affecting music memorisation both directly, as it is actively involved in the memorisation process, and indirectly through sight-reading. Sight-reading is one of the most prominent musical activities to be affected by working memory capacity and processes (Aiello and Williamon, 2002), as it requires rapid on-line processing of musical material. The importance of sight-reading in memorisation is supported by Chaffin et al (2003) who argue that the initial 'artistic image' of the piece, formulated during sight-reading, serves as a starting point for the shaping of the final performance of the piece and effectively leads to the desirable memorised performance. This view is also empirically supported by many music teachers and conductors, who advocate the need for an efficient first sight-reading of any given piece in order to be able to construct a solid representation of the music, which will be subsequently

elaborated on. On the other hand, when music memorisation is tested in experimental conditions, a considerable load is directly placed on working memory as there is a restricted timeframe during which subjects are called to perform the memorisation task; researchers have repeatedly associated the continuous mental and physical engagement with the piece, involved in the performance of memorisation tasks, with the activation of short-term processes and stores (Williams, 1975; Palmer, 2005). The importance of these processes and the connection between initial perception, encoding and short-term storage has been explained previously (see sections 1.2 and 1.3).

The ability to memorise proficiently has long been regarded as an invaluable asset for musicians and a skill well worth cultivating; even as early as 1915, Hughes recognises the growing demand for, and purported superiority of, soloists who perform music from memory (Hughes 1915). Williamon (1999, 2002) has investigated experimentally the response of audiences towards memorised performances and has shown that both musicians and non-musicians value memorised performances more than performances with the score, in terms of musical understanding, communicative ability, technical proficiency and overall performance quality. Extraordinary memory also seems to be a key feature both in cases of musical prodigy, where it can compensate for lack of practice time (Ruthsatz and Detterman 2003), and in musical savants (Heaton, 2003; Sloboda, 2005).

One of the main focus points for research has been the identification of the distinct modes and techniques musicians use in order to memorise music; various studies have examined the processes and mechanisms involved in the memorisation of music by instrumentalists and vocalists (Chaffin and Imreh, 1997; Ginsborg, 2002), and possible methods of refining memorisation skills (Hughes, 1915; Matthay, 1926; Ginsborg, 2002; Ginsborg and Sloboda, 2007). In the largest part of this body of research, four basic types of music memory are identified as aural, visual, kinaesthetic and conceptual memory, which reflect the four basic strategies for memorising music: memorisation based on 'how the music sounds' (pitch, timbral and rhythmic information, pitch contour); memorisation based on the visual image of the printed music; memorisation based on bodily activity when performing the piece; and memorisation based on knowledge of the music's structural characteristics. None of the above, however, can account for the kind of memorisation taking place when using solfège. Solfège can

function as the basis of a fifth type of musical memory, which is quasi-verbal, as opposed to the verbalised nature of conceptual memory and is based on pitch names: **pitch label memory**. The notion of pitch label musical memory is not new, although it has not been identified as a separate function per se; Hughes (1915) identifies three types of musical memory in pianists: auditory, visual and 'finger' memory, but in addition refers to the ability to 'say the notes' in connection to either of the first two memory types; the ability to 'say the notes' is also reported by Sloboda (1985) as providing a possible advantage in performing memorisation and aural training tasks. On the other hand, the practice of singing the music to be memorised can also be regarded as part of 'physical', i.e. kinaesthetic practice (Lim and Lippman, 1991); however, while the proposed pitch label memory can use the kinaesthetic mode for its output, it is also completely functional without extrovert expression or rehearsal.

Apart from the distinction between different types of musical memory, two other crucial functions in the memorisation process are the creation of appropriate performance cues and retrieval practice. Chaffin and Imreh (1997) identify three types of performance cues: basic, interpretive and expressive; the latter two regard issues of tempo, phrasing, dynamics and mood, while 'basic' cues concern more 'objective' characteristics of the music, such as conceptual chunks, fingering and general technical difficulties. Solfège helps in the formation of a new set of 'basic' cues, based on the retrieval of the verbal trace of certain pitch sequences. On the other hand, retrieval practice has been shown to be a priority for proficient musicians, even in very early stages of the memorisation process (Chaffin 2007); efficient retrieval is also heavily dependant on the formation of appropriate cues. This presumably means that the potential solfège provides for rapid and relatively effortless formation of cues is an extremely important aid for quick and efficient memorisation.

Most of the above findings refer mainly to the memorisation of tonal music; research on memorisation of atonal pieces is more scarce and usually refers to short atonal sequences. Imberty (1993) discusses and reviews literature on the perception of atonal music, referring to properties of atonal music which can lead to better cueing, recognition and recall. Given his focus on proposing possible ways to conceptualise atonal music meaningfully and effectively, all mnemonic features in Imberty's discussion fall into the category of 'conceptual' memory; he concludes that

conceptualisation in atonal music is much more complex an issue than in tonal music. This is assumed to entail the need for all theories of atonal music to allow for much 'psychological elaboration' from the part of the listener in order to formulate a perceptual image of the piece. If this is the case, though, then obviously the task of the musicians memorising atonal music becomes infinitely harder, as the tools they have readily available in the case of tonal music cannot be adapted to fit atonal patterns – or non-patterns; in this light, Imberty's proposal is actually depriving the musician of the solidity of the most potent memorising tool: conceptual memory. Since, thus, atonal music is much more reliant on intervallic relations than chords (Imberty, 1993) and given that solfège enhances understanding and ease in interval perception and usage, it is possible that solfège can prove an invaluable tool, especially for the memorisation of atonal music.

In sum, the potential benefits solfège can provide towards more efficient and easy music memorisation are the following:

- Solfège generates an additional encoding, which is important both at the level of perception (spread of encoding) and at the level of memorisation, as it creates additional associations with the music to be memorised.

- The additional encoding generated by solfège is also important in the sense of quality of encoding: it creates a verbal memory trace for the music itself, as opposed to meta-musical characteristics; this 'pitch label' memory generates a distinct set of cues.

- The use of solfège provides the opportunity for additional practice, which has been shown to be crucial to the memorisation outcome (Williamon, 1999).

Chapter 2

Methodology

2.0. Introduction

Having reviewed the relevant research on solfège, perception, memory and music memorisation, certain omissions have become apparent: although solfège fulfills key criteria for potentially playing a major role in music cognition, there is hardly any reference to it in the literature. The purpose of the present study is to provide an initial understanding of the complex phenomenon that is solfège, by addressing three key areas:

1. The impact of perceiving music through solfège syllables, as expressed in the performance of memorised music.
2. The impact of using solfège syllables in the performance of memorised music.
3. The dual linguistic–musical nature of solfège syllables.

The first two areas, regarding the impact of solfège syllables on music perception and memorisation, are going to be addressed in two closely related, observational experiments: the first experiment will provide a comparative examination of the output of music memorisation produced by musicians who know solfège (hereon solfège musicians) and musicians who have never been taught solfège (hereon non-solfège musicians). The second area is best explored by a longitudinal, within-subjects design: Experiment 2 is going to look at differences in the music memorisation output by non-solfège musicians, before and after they learn solfège. The nature of solfège syllables is going to be investigated in a third experiment, comprising a set of shorter experiments testing the extent of interference between solfège syllables, musical and verbal material.

The empirical approach adopted in each one of the three experiments and the specifications of their design were selected after careful consideration of issues regarding ecological validity of the research; although a rigorous methodological framework was desirable, maximum effort was made to avoid the use of a reductionist paradigm, which would be detached from real-life music and would compromise the potential to generalise research findings. The goal of the first two experiments was to study the effects of solfège in a controlled environment: short musical pieces were specifically created for the purposes of the particular research, participants had a strict timeframe within which they should complete the tasks and they were also given specific instructions regarding the performance of the tasks. On the other hand,

participants acted in the context of 'natural purposeful activity' (Neisser, 1976, p. 7), as they were not asked to memorise isolated pitches or short, nonsensical pitch sequences; instead, set pieces for memorisation were 8 bars long, resembling 'real' music passages. Moreover, musicians were observed and recorded on the set tasks in their natural environment: the practice room, thus offering a naturalistic framework for the experiment (Bahrick, 1996).

The set of experiments comprising Experiment 3, on the other hand, were all typical laboratory experiments, loosely following the paradigm set by research investigating the effect of unattended music on reading comprehension, verbal and phonological tasks (Henderson, Crews and Barlow, 1945; Martin, Wogalter and Forlano 1988; Salame and Baddeley, 1989). The selection of such a strictly defined design was necessary in this case, as the object of investigation was not a large-scale process, as was the memorisation of a music excerpt in Experiments 1 and 2, but rather the interaction between elemental units of musical and verbal memory; this objective called for a targeted, controlled procedure in order to test fine-tuned elements of memory and yield results that could be generalised beyond the sample. The structure of the experimental procedure followed the model of working memory span tests and especially the Automated operation span task, devised by Unsworth et al (2005).

2.0.1. Aims and objectives

In order to examine the three areas of interest described in the previous section, five research aims and their corresponding research questions were defined:

Aim 1: To examine the effects of solfège use on a music memorisation task performed by adult musicians.
Research question 1: Does the early learning and use of solfège note names affect music memorisation performance in adulthood?

This question was tackled in Experiment 1, in a comparative, cross-cultural study looking at the performance of solfège musicians in relation to that of non-solfège musicians. The two types of musicians came from different countries: solfège musicians

came from countries that incorporate solfège teaching since the early stages of their music education system, whereas non-solfège musicians came from countries where solfège is not found in the core of music studies, at any level. This inevitably raised the issue of cultural factors playing a role in the performance of memorisation tasks (Robertson, 1998, in Madsen and Madsen, 2002, p. 113); in order to balance potential differences resulting from a difference in approach to music education other than the use or not of solfège, several restrictions were applied in the choice of sample and in the construction of musical stimuli:

- **The sample** in this experiment consisted exclusively of university music students, with specifications as to the minimum number of years of instrumental lessons they have had prior to the tests; as research has shown that age and musical experience play an important role on the performance of recall tasks (Halpern, Bartlett and Dowling, 1995), participant selection methods in Experiment 1 intended to achieve maximum homogeneity of age, experience and training across the sample.

- **The stimuli** on which the subjects were tested were constructed so as to pose minimal technical challenges to subjects; on the other hand, the pieces had an increased degree of memorisation complexity. The creation of the pieces followed the criteria set in the work of Madsen and Madsen (2002) regarding metre and tonality.

Aim 2: To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians.

Research question 2: Does the early learning and use of solfège note names affect long-term retention of memorised music in adulthood?

This question was also tackled in Experiment 1, as the design was extended to include long-term recall; the term long-term in this case implies the use of long-term memory stores, rather than a long duration. Solfège and non-solfège performance of a memorised piece was compared 24 hours after the stimulus presentation.

Aim 3: To examine the effects of solfège use on a music memorisation task performed by adult musicians who have been taught solfège recently.

Research question 3: Does the use of solfège note names affect music memorisation performance when solfège has been learnt during adulthood?

As fixed-do solfège is normally taught during childhood, along with the first music-theory lessons, Experiment 2 explored the possibility and utility of teaching solfège to adult musicians. The effects of solfège learning on music memorisation were examined by looking at the performances of students before and after they had learnt solfège.

Aim 4: To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians who have been taught solfège recently.

Research question 4: Does the use of solfège note names affect long-term retention of memorised music when solfège has been learnt during adulthood?

As in the case of Aim 2, the design of Experiment 2 was extended to compare the long-term recall of a melody by the same students, before and after they had been taught solfège.

Aim 5: To examine the differences between subjects using solfège syllables and subjects using letter names, regarding the extent and nature of interference between verbal and musical tasks being performed simultaneously.

Research question 5: Is solfège different from other note-reading systems, in terms of the interaction taking place between note names and linguistic elements?

An answer to this question was attempted to be given by a series of short, computerised experiments, comprising Experiment 3. In these experiments, the performance of solfège and non-solfège subjects was compared in tasks testing the interactions between pitch labels, whether solfège syllables or letter names, and verbal elements in the phonemic, morphemic and lexical level.

2.1. Solfège lessons

For the purposes of this research, a group of participants, who were music students in the University of Hull, followed solfège lessons taught by the researcher in the Department of Drama and Music of the University of Hull. The first series of lessons took place between February and May 2009 and were offered only to 7 students, who were participating in the pilot study and constituted the experimental group. From September 2009, however, solfège lessons were offered in the Department as a non-compulsory aural training programme. This training was open to all students and, those who attended the required amount of lessons were able to claim assignment credit for performance-related modules as part of their degree programme; at the end of the year, the positive feedback received by all participating students allowed solfège lessons to continue during the next academic year, independently of the purposes of the present research. In the subsequent sections, the evolution of the solfège lessons from the pilot study to the final training programme will be described.

2.1.1. Solfège lessons in the pilot study

As the pilot study needed to take place within a restricted timeframe, there was only time for 8 solfège lessons between February and May 2009. The lessons lasted 50 minutes each and took place on a weekly basis, in a large office equipped with a Yamaha Clavinova and a whiteboard. The room was set up for the lessons so that students were sat next to each other in one row of tables facing the teacher, the clavinova and the whiteboard.

2.1.1.1. Teacher

At the time of the pilot study, the researcher had never taught solfège before; her only experience with solfège lessons was as a student, for 7 years in a Conservatoire (from 5 to 12 years old) and for 2 semesters in the University, as a postgraduate student. Solfège teachers, in countries that use solfège, usually fall into two main categories: the first is

the most common one, in which solfège teachers are required to teach children aged between 4 and 15 how to read and sing music; these teachers can either be Music graduates from a university, or holders of a Degree in Music Theory, or, in fewer cases, professional singers, pianists, or instrumentalists of another specialty. The second category is solfège teachers in Universities or other higher education institutions; these teachers usually have completed postgraduate studies in music on a performance-related subject and train students in advanced aural skills, dictation, singing complex rhythms and polyphony.

In the present study, the teacher was a professional cellist, with a Masters degree in performance. Her only singing experience was as a member of a youth choir for 8 years, participating in international festivals and competitions; she also had experience teaching adults as a director of an amateur choir for two years.

2.1.1.2. Resources

There are countless solfège books used by teachers all over the world to teach fixed do solfège; perhaps the most widespread and well-known are the 'Lemoine method' (Dannhäuser, Lavignac and Lemoine, 1891) – named after the French publishing house issuing the volumes; the actual title of the books is *Solfège des Solfèges* – the 'Dandelot' (Dandelot, 1928) and Kodály's method (Kodály, 1941). The problem with most solfège books, however, is that they are principally targeted at children or, at least, at adults who do not know how to read music. It is very rare that adult music students, who have already attained a high level in music performance or theory using letter names or numbers, would go into the trouble of 're-learning' music reading using solfège syllables. For this reason, the vast majority of solfège books are addressed to children or amateur musicians and thus begin with extremely simple melodies in treble clef, using simple rhythms with big rhythmic values and stepwise movements with a few thirds, fourths or fifths; the progress onto more interesting and complex musical material is very gradual and slow, as solfège instruction is normally an integral part of music theory lessons, undertaken during a long period of time. In the case of solfège instruction in higher education, teachers rarely make use of a single book; the musical

material studied is usually 'real' music, either directly written for voice or choir, or transcriptions of pieces for instrumental ensembles for voices.

The fact that the students taking part in the pilot study were all musicians with Grade 8 or higher in their respective instruments, as well as the limited time in which solfège instruction had to take place, rendered the exclusive use of one of the available teaching manuals highly inappropriate. Moreover, as the seven students were participating in solfège lessons on a voluntary basis, they had to be kept motivated and interested, but at the same time not overwhelmed by the difficulty of the task. Bearing the above in mind, the model of solfège lessons in higher education institutions was followed: a corpus of 70 short pieces was put together, using music exercises and songs from various sources. The material was carefully selected so it was as age- and level-appropriate as possible; the structure of the handbook was the following (an enhanced version of the handbook that was used in the main study can be found in Appendix 1):

- **Introductory exercises** (exercises 1–7). Acquaintance with solfège syllables and simple exercises using notes C to G in treble clef.
- **Preliminary exercises** (exercises 8–12). Simple singing exercises using all 7 notes.
- **Warm-up** (exercises 13–15). Standard warm-up exercises used by singers and choirs; these were provided in written form, so that students would have the visual aid which would help trigger the use of the correct solfège syllable.
- **Main core** (exercises 16–47). Pieces in treble clef, with an increasing level of difficulty.
- **Bass clef** (exercises 48–54). 7 exercises in bass clef with an increasing level of difficulty.
- **Two-voice exercises** (exercises 55–70). Two-voice pieces with both voices in treble clef.

The exercises in the handbook were taken from the following sources and collections:

A. Dannhäuser, A. Lavignac and H. Lemoine, *Solfège des Solfeges* (Milwaukee: G. Schirmer, 1891)

Z. Kodály, Tizenöt Kétszólamú Énekgyakorlat (Budapest: Editio Musica, 1941)

N. Ladukin, Monophonic Solfège (Moscow: Editio Musica, 1980)

In addition to the pieces in the handbook, on certain occasions students were given photocopies of choral and orchestral music which they were performing at the time in one of the University ensembles.

2.1.1.3. Lessons

After the introduction of solfège syllables in the first lesson, students were asked to always use them when singing or referring to pitches during the lesson. Regarding the pronunciation of the solfège names, no strict rules were followed: the teacher explained that she would always be using 'do, re, mi, fa, sol, la, si' due to long-term experience with these names, but the students were free to use variants of the syllables, such as 'ray' instead of 're', 'so' instead of 'sol' and 'ti' instead of 'si'; they were, however, advised to consistently use the syllables they chose at the beginning rather than constantly switching between alternatives.

Due to the uncertainty caused by the uniqueness of the venture of teaching solfège to adult musicians, there was not a strict macroscopic plan for the lessons, with specific learning aims to be achieved in each session. Even the handbook had been compiled to cater for many possible outcomes: the two-voice exercises provided a musically interesting alternative of the easy-to-intermediate one-voice exercises in treble clef, in case solfège proved too hard to learn in such a short time; the more advanced exercises and the pieces in bass clef would be used in case the students felt they wanted to be more challenged. The general aim was that, by the end of the semester, the students would be able to sing at least intermediate level pieces using solfège syllables with relative ease; an example of the desired level at the end of the lessons is provided in Music Example 2.1:

B (57)

Music Example 2.1: Exercise no. 34 from the Solfège Handbook in the pilot study, taken from Nikolai Ladukin's *Monophonic solfège* (1980).

Microscopically, however, the lesson had a precise and predetermined structure, which was followed, with small deviations, in every session:

1. Warm-up (5 minutes). This consisted of simply singing scales in solfège in the first lessons; once the students started feeling more familiar with solfège names, the dedicated warm-up exercises from the handbook were used.

2. Intervals (10–15 minutes). As the aim was to get familiar with solfège syllables as soon as possible and considering that some of the students were not accustomed to singing in general, it was thought best to begin teaching solfège names focusing on their fixed-do environment and function. In each lesson, an interval was presented and elaborated on: students learned to sing it in its different variants (minor, third, perfect, augmented, diminished), either reading it, or asked to form it starting from a given pitch, or recognizing it when played or sung to them; in all these instances, students were asked to use solfège names whether singing or simply referring to a note. Intervals were studied in the following order:

   - 1st lesson: minor and major second
   - 2nd lesson: minor and major third
   - 3rd lesson: perfect fourth and fifth
   - 4th lesson: augmented fourth/ diminished fifth
   - 5th lesson: minor and major sixth
   - 6th lesson: minor and major seventh
   - 7th lesson: revision of all intervals

3. Singing exercises in unison (15–20 minutes). All students were asked to sing together exercises in treble clef from the handbook.

4. Singing two-voice exercises in pairs or all together (10–15 minutes). Students sang either in pairs or were divided in two groups; in most cases after the first lesson, it was the students who chose the two-voice exercises they wanted to sing from the handbook.

In terms of singing technique, only the basic choral instructions were given: students were asked to sit in an upright position, sing with an open mouth and enunciate. Students were also encouraged to count using their right hand, as if conducting themselves; most students, however, found this quite hard. In the first two lessons, students were very self-conscious and hesitant about singing in front of the other students; apart from the only singer in the class and one student who said was used to singing with others from music lessons in her school years, all the others had expressed doubts regarding their ability to sing. As the lessons progressed, they overcame their initial reservations and, by May, they all were considerably more confident about singing and sight-singing in front of the class.

The solfège syllables themselves posed an additional obstacle: although in the beginning they provided a helpful common base for singing, the more the students improved in sight-singing, the more they were inclined to abandon solfège names and sing what they knew was correct, using a neutral syllable (like 'na' or 'ta'). All students, however, noted that this was due to them not practising solfège at home; as they were all in the last semester of their final year and solfège was a voluntary programme, they did not invest any more time or effort in it than what they spent coming to class. This was reported both in private communications, in which all students admitted they felt solfège would have been more helpful if they had more time to work on it, and in the anonymous feedback forms they were asked to complete at the end of the lessons (Appendix 1): out of 7 students, all 3 who answered 'No' to the question 'Do you feel you have learnt the material covered in the lessons sufficiently?', responded in the next question that this was due to 'Lack of individual practise'.

2.1.1.4. Learning outcomes and exam

By the end of the semester, all students were able to use solfège efficiently in pre-learnt pieces, regardless of difficulty, and in sight-singing simple pieces. Their aural skills and their general ability to sight-sing improved substantially during the semester; this, however, could not be attributed exclusively to solfège, as the simple fact that the students received additional aural classes was bound to have an impact on their aural abilities.

At the end of the course, an informal exam took place, in order to screen participants who would be able to use solfège in the memorisation experiment for the pilot study. The exam was split in three parts:

1. Singing a piece they had prepared from the handbook. The students were given a selection of 3 pieces from the handbook, from which they had to choose and prepare 2 for the exam.

2. Sight-singing an easy piece.

3. Singing a two-voice piece they had prepared from the handbook with another student. The students were given a selection of 3 pieces from the handbook, from which they had to choose and prepare one for the exam.

All three exam tasks had to be performed using solfège. The students came in the exam in pairs; as there were 7 students in the class, one of them appeared twice to cover the second-voice of the 7th student's exam, but was only evaluated for his original performance. Each of the three tasks received a mark between 1 and 10; at the end of the exam, students who had scored more than 15 points were asked to perform the pilot memorisation experiment using solfège. Out of the seven students, five scored more than 15 points; one of them, however, suffered an injury and was not able to attend the experimental sessions in May, so ultimately only four students were asked to use solfège in the experiment.

2.1.1.5. Questionnaires and feedback

At the end of the solfège course, the students were asked to complete a short questionnaire as well as an anonymous feedback form, regarding their experience with solfège. In both forms, when asked to rate the difficulty of the solfège lessons overall, 4 students rated solfège lessons as being 'Difficult' and 3 students as 'Neither difficult nor easy'. In the feedback form, when asked whether they felt they had enough time to learn solfège adequately, 3 students answered 'Yes' and 4 students 'No'; 2 of the students who had answered no thought they would require 2-4 more weeks of lessons and two students required 1 semester more of solfège lessons in order to learn it more adequately. All the above responses, combined with exam results, led to the conclusion that more time had to be devoted to solfège lessons for the main study.

Regarding the contribution of solfège in the students' overall understanding of music, 6 out of the 7 students answered in the feedback forms that solfège had improved their musicianship; the fields in which they felt solfège contributed are presented in Table 2.1:

Which aspect of your personal music practice/performance did solfège improve?

	Aural skills	Sight-reading	Interval recognition	Practice technique
Number of responses	4	3	2	1
Participant's instrument	V, S, W, P	V, P, P	S, W	W

V = voice, S = string instrument, W = wind instrument, P = piano

Table 2.1: Students' responses regarding the aspect of musical performance they felt had benefited from solfège lessons.

Once more, it should be noted that the improvement on specific, aural-related skills should not be attributed to solfège per se, but rather to the fact that these students had received tailored aural lessons in addition to the modules in their degree programme.

Nevertheless, the positive feedback received for the pilot lessons supplied a powerful argument that motivated students to participate in lessons taught for the purposes of the main study.

2.1.2. Solfège lessons in the main study

In order to attract as many participants as possible, the solfège course for the main study was advertised to all music students, regardless of their suitability to complete the memorisation tasks of the experiment itself. Motivated by the previous years' students' positive feedback and by the fact that solfège attendance enabled them to claim assignment credit for certain performance modules, a total of 31 students signed up for the solfège programme; out of those, 18 students were also participants in the experiment, forming the experimental group. The remaining 13 students had decided to take up the lessons in order to develop their aural skills and did not take part in the memorisation experiments; this group will hereon be referred to as the 'aural class'. Because of the large number of participating students, the solfège programme was divided in two groups, taught in different times during the same day; participants in the experiment were mixed with students from the aural class in these groups. The lessons were weekly, lasted 60 minutes and were offered between October 2009 and May 2010, excluding holidays and examination periods; a total of 19 lessons were taught during the academic year, 9 during the first and 10 during the second semester. At the beginning of the second semester, 4 students from the experimental group and 4 students from the aural class dropped out of the programme, mainly because of timetable clashes with other modules or because they felt they could not invest the time required for the lessons; a total of 23 students completed the solfège programme. A minimum attendance of 15 lessons had been set as a prerequisite both for participation in the experimental group and for claiming assignment credit; at the end of the programme, 14 students from the experimental group and all 9 students from the aural class had fulfilled this criterion.

Lessons took place in a seminar room equipped with an upright piano and a whiteboard; the same setup as in the pilot lessons was used for the main study lessons, with the students sat in rows of tables facing the teacher.

2.1.2.1. Resources

Participants in the pilot study had given positive feedback regarding the resources and the teaching method used in the solfège lessons; the main conclusions from the pilot solfège lessons, drawn from students' comments and from the teacher's observations in class and in the exam were the following:

- It was relatively easy to learn the solfège note names; extended practice, however, was vital for the efficient use of the syllables in singing. For this reason, the introductory, basic-level singing exercises were kept to a minimum and more pieces were added in the main core of the handbook.

- Two-voice exercises were the most popular in the handbook. More two-voice, as well as some three-voice pieces were added to this end.

- Students expressed a preference in bass clef exercises, as these presented a bigger challenge to them; they had also expressed the desire to learn to sing in more clefs, if time allowed. In order to accommodate these preferences, alto clef pieces and pieces with clef changes were added in the handbook.

- The study of intervals was reported by all students to be 'extremely interesting and useful'; this section of the lesson had been conducted impromptu in the pilot study, without the use of any textbook resources. The basic core of teaching intervals was kept the same, but the teacher also used Lars Edlund's *Modus Novus* (1963) in the main study, for the teaching of intervals and dictation outside a tonal context, so as to promote the study of intervals to a more advanced level.

Bearing these points in mind, the solfège handbook used in the pilot study was amended and enhanced to fit the purposes of the main study (the full handbook can be found in Appendix 1); the structure of the final handbook was the following:

- **Introductory exercises** (exercises 1–5). Acquaintance with solfège syllables and simple exercises using notes C to G in treble clef.

- **Preliminary exercises** (exercises 6–12). Simple singing exercises using all 7 notes.

- **Warm-up** (exercises 13–15). Standard warm-up exercises used by singers and choirs; these were provided in written form, so that students would have the visual aid which would help trigger the use of the correct solfège syllable.

- **Main core** (exercises 16–46). Pieces in treble clef, with an increasing level of difficulty.

- **Bass clef** (exercises 47–56). Exercises in bass clef with an increasing level of difficulty.

- **Exercises with clef changes** (exercises 57–62). Exercises with changes between treble and bass clefs.

- **Alto clef** (exercises 63–69). Exercises in alto clef.

- **Exercises with clef changes** (exercises 70–75). Exercises with changes between treble, bass and alto clefs.

- **Two and three-voice exercises** (exercises 76–101). Exercises with two or three voices; all voices in treble clef.

Exercises in the handbook of the main study were taken from the same Lemoine, Kodály and Ladukin textbooks as in the pilot study. In addition, students were given handouts of transcriptions from Marie-Jeanne Bourdeaux's *25 leçons progressives* (1975), in which staff notes were replaced by rhythmic values on a single line, to be used for rhythm practice.

2.1.2.2. Lessons

The structure of the lessons in the main study was modelled on the lessons in the pilot; certain changes, however, were made in order to accommodate students who were more interested in the aural rather than the solfège part of the lessons. As the two groups included students from both the experimental and the aural groups, the lesson was structured to resemble more an aural training class rather than a solfège class run for purely experimental purposes. This may seem, at first glance, as a compromise in relation to the research aims; in practice, however, solfège lessons in the main study were more successful than in the pilot for two reasons:

1. The fact that there were more lessons, with more time between them, made solfège incorporation easier and more natural; this is in accordance with the Total Time Hypothesis and the Distribution of Practice Effect principles (Baddeley, 1999), which suggest that the more time is spent on a task, and the more this time is broken down in smaller units rather than spent in one long session, the more efficient learning becomes. In the lessons of the main study, solfège was omnipresent as the medium of communication between teacher and students; as such, the fact that it was not being constantly practised as an end to itself allowed students to use solfège creatively and productively in their aural training, appreciating its function in its natural environment of a music theory/ aural training class.

2. The feeling of accomplishment students gained in the process of refining their aural skills served to maintain their motivation to stay in solfège classes. Given the duration of the course, students would have most probably lost interest if they felt they were just participating in a research project that had little to offer them as musicians; the sense of accomplishment also helped tremendously with the rapport between teacher and students.

In summary, the structure of the lessons in the main study was the following:

1. **Warm up** (5 minutes), was kept the same as in the pilot study.

2. **Intervals** (10–15 minutes). The method of interval instruction was kept the same as in the pilot study, with the addition of some listening/ singing exercises from Edlund's *Modus Novus*. The overall plan was slightly changed, according to the new timeframe:

   1st lesson: Minor and major second

   2nd lesson: Minor and major third

   3rd lesson: Minor and major sixth

   4th lesson: Juxtaposition of thirds and sixths

   5th lesson: Perfect fourth and perfect fifth

   6th lesson: Tritones

   7th lesson: Minor and major seventh

   8th lesson: Juxtaposition of seconds and sevenths

   9th lesson (end of 1st semester): Intervals revision.

In the second semester, the time used for the intervals was dedicated to element no. 4 (dictation).

3. **Singing in unison or in groups** (15–20 minutes), was conducted using either one or two-voice exercises from the handbook.

4. **Dictation** (20 minutes). This was practised in the second semester, as a natural passage from forming specific intervals upon request and interval identification. The dictation mostly used melodic and atonal pieces, taken from Edlund's *Modus Novus*; tonal dictation was done in a few occasions using unfamiliar pieces from the handbook. At the end of each dictation session, students were asked to sing the corrected piece, using solfège. Harmonic dictation was also practised to some extent, where minor/major and diminished/augmented chords were studied in separate lessons; in many occasions, pitch clusters from *Modus Novus* were also used.

5. **Rhythm/ Singing and playing simultaneously** (10–15 minutes). At the beginning of the course, students had been given a brief description of what the aural classes would entail and were encouraged to voice their opinion if they wanted anything else added to the material. At that point, many students expressed the desire to focus on rhythm as well; this was noted and, in most lessons, some time was dedicated specifically to rhythm, using transcriptions from M.J. Bourdeaux's *25 leçons progressives* (1975). When this did not happen, students used the time to practise two-voice pieces by singing one voice while playing the other on the piano.

2.1.2.3. Learning outcomes and exam

By the end of the solfège course, all students were able to sing a prepared piece efficiently and to sight-sing pieces using solfège. As in the pilot study, most students in the main study had little experience with singing, apart from choir participation, so many of them reported at the end of the course that they had linked sight-singing with solfège syllables.

The solfège exam took place in May 2010 and comprised four different elements:

1. **Singing a piece in treble clef.** Students were asked to prepare four pieces from the handbook and performed one in the exam.

2. **Singing/ playing a piece with clef changes.** Students were asked to prepare two pieces with clef changes and perform one of them in the exam. Participants from the experimental group were asked to perform this task by singing, but students from the aural class were asked to perform it on their 1st study instrument.

3. **Playing/ singing back a 4-bar melody.** This was a form of melodic dictation, but instead of the students having to write the melody, they had to either sing it back using solfège names (experimental group), or play it on their 1st study instrument (aural group). In both cases, students were only allowed to correct each note once, but they could take as much time as they wanted before they started playing/ singing back the piece.

4. **Play and sing or sing in pairs.** Experimental group students were asked to prepare a two-voice exercise and sing it while playing the other voice on the piano. Aural class students were given a choice between this task and preparing 2 two-voice pieces to sing in pairs.

Each task was given a maximum of 20 points; in addition, there was an 'attendance bonus', which was worth 20 points if someone had attended all 19 lessons, with 2 points deducted for each missed lesson. If the overall mark was above 50, it enabled the student to claim assignment credit and to participate in the experiment as a solfège participant. All students who participated in the exam managed to score above 50 points; there were 2 students, both from the experimental group, who did not attend the exams although they had fulfilled attendance demands. Overall marks ranged from 52 to 97; the mean score was 74.62 (Table 2.2):

Statistics

Total Mark

N	Valid	21
	Missing	2
Mean		74.62
Std. Error of Mean		2.732
Median		75.00
Std. Deviation		12.520
Variance		156.748
Range		45
Minimum		52
Maximum		97

Table 2.2: Descriptive statistics of marks scored in the solfège exam.

2.1.2.4. Questionnaires and feedback

As in the pilot study, all 23 students attending solfège lessons were asked to complete a questionnaire and an anonymous feedback form regarding their experience with solfège; a total of 22 questionnaires and 20 feedback forms were received back completed.

In general, students' responses both in the feedback forms and in the questionnaire show that they had found the lessons mostly difficult or average (Figure 2.1); feedback forms also revealed that most students assessed their own solfège skills at that point to be average or good (Figure 2.2):
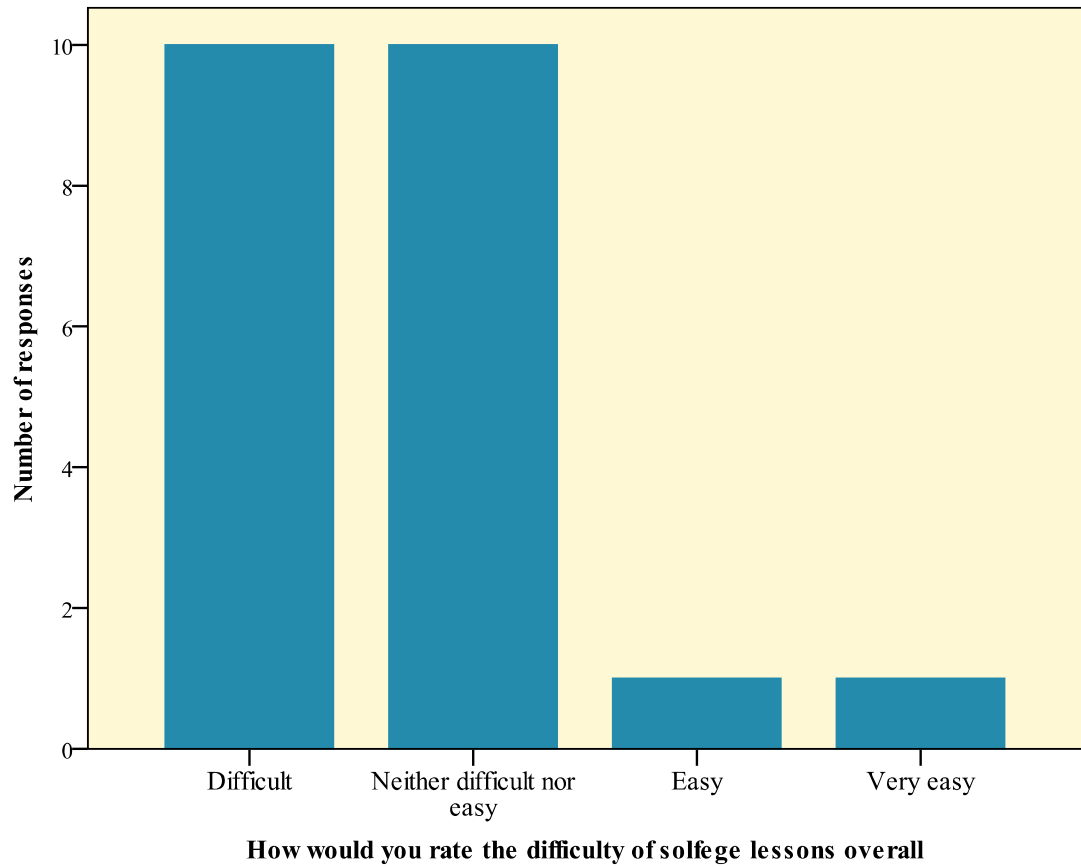
Figure 2.1: Questionnaire responses regarding the difficulty of solfège lessons in the main study. There were five possible answers, ranging from 'Very difficult' to 'Very easy'; none of the students selected the 'Very difficult' option.
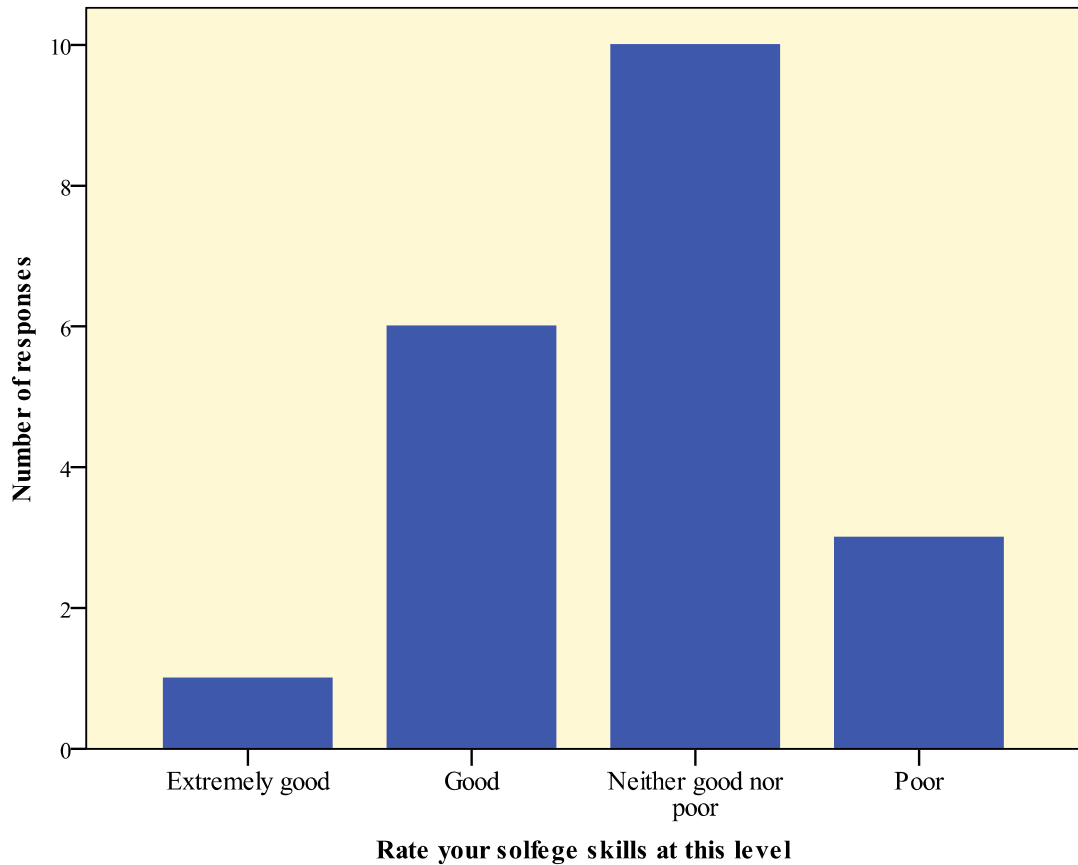
Figure 2.2: Feedback form responses regarding the self-rating of students' solfège skills at the end of the lessons in the main study. There were five possible answers, ranging from 'Extremely good' to 'Extremely poor'; none of the students selected the 'Extremely poor' option.

Most students (N=17) in the main study found that the time they had to learn solfège was adequate, while only three thought that they would require more time to learn it; these students answered they needed up to one more semester of lessons in order to learn it adequately. 19 out of 22 students responded that they felt they knew solfège well enough to use it when learning an unfamiliar piece of music; out of them, 15 answered they would use solfège when learning unfamiliar music.

Students' attitude towards lessons was generally very similar in the main and the pilot study; in the main study, all 7 students who thought they had not learnt the material covered in the lessons sufficiently, also responded that this was due to lack of individual practice. Both in the questionnaires and in the anonymous feedback forms, all students answered that solfège skills had contributed in their musical skills in general; it is again the case, however, that the improvement of musical skills cannot be attributed to solfège

per se as students were likely to have improved regardless of using solfège in their aural training course.

The only occasion in which solfège could potentially be credited with the general improvement of musical ability, was the case of LK, a first-year undergraduate from the aural group, who had come to solfège lessons without almost any prior knowledge of music theory. He was a skilled jazz-guitar player and was familiar with letter names and could read from a score, although he normally performed without written music; during the solfège course he was attending the aural training module for first-year music students as well as private music theory lessons. He reported, both in the feedback forms and in private communications, that solfège had helped him more than the letter system used in the aural training module and in the private lessons, because he felt it had provided him with 'a point of reference' regarding written music. He also reported that, although he was familiar with letter names before, practising solfège singing had made him link the notes written on a staff with the musical sound of their solfège name; this, he reported, was very different to knowing the letter names because solfège syllables felt more musically relevant, since they were always sung. It is important to note at this point that the particular student was not aware of the background, aims and hypothesis of the present research.

Regarding the incorporation of solfège in a University music curriculum, the vast majority of students answered that they approved; the responses were very similar to the question of solfège incorporation in a school curriculum (Figure 2.3, Figure 2.4):

How would you feel about solfege instruction being incorporated in a University curriculum?
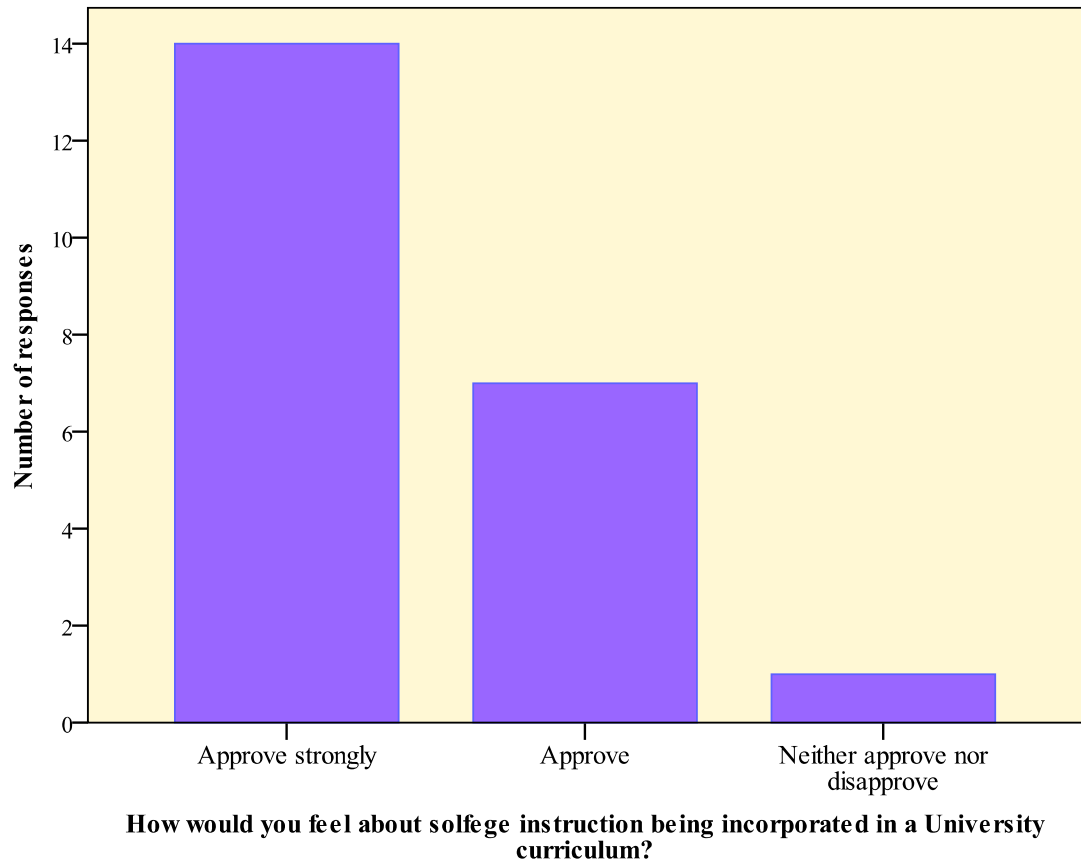
Figure 2.3: Questionnaire responses regarding the potential incorporation of solfège instruction in a University music curriculum. There were five possible answers, ranging from 'Approve strongly' to 'Disapprove strongly'; none of the students the 'Disapprove' options.

How would you feel about solfège instruction being incorporated in a school curriculum?
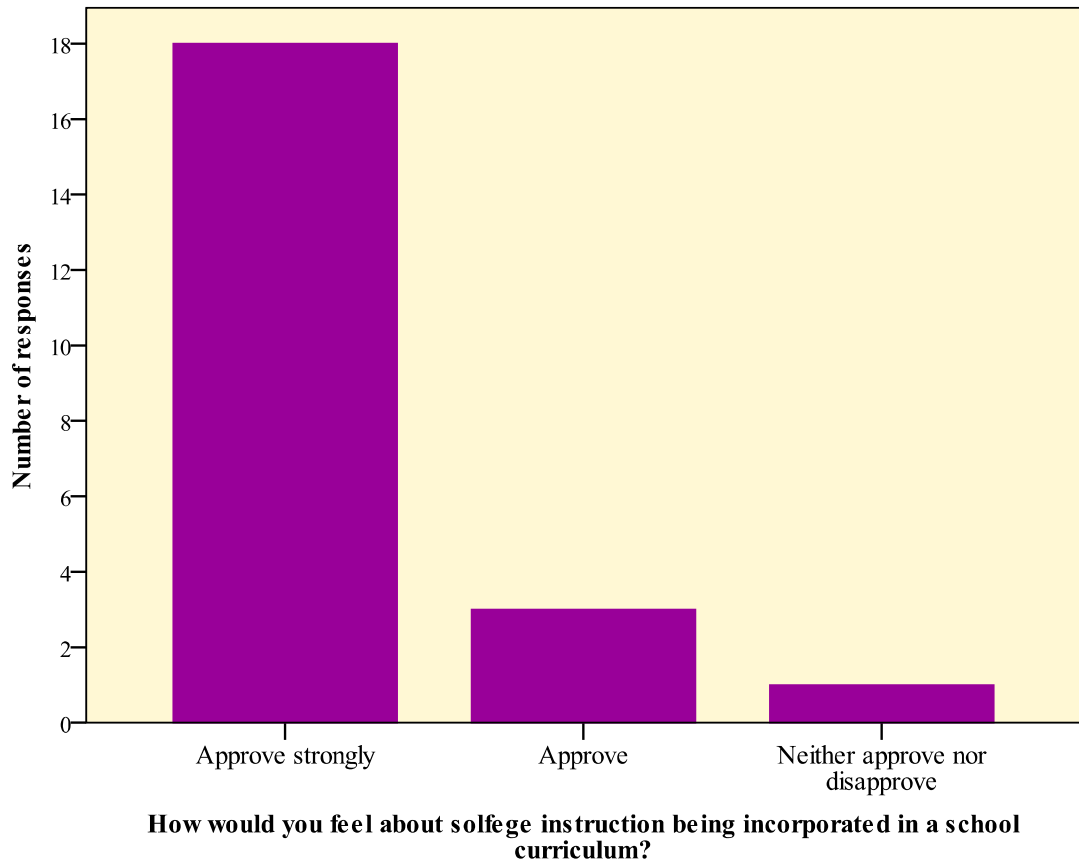
Figure 2.4: Questionnaire responses regarding the potential incorporation of solfège instruction in a school music curriculum.

2.1.3. General remarks regarding solfège lessons

The reception of solfège lessons overall was very enthusiastic, especially in the main study; many students commented they would like to see solfège become a compulsory module in their music degree. It was not within the scope of the present study to determine which aspect of solfège instruction in specific was the one that made it so popular with students; general student feedback, however, indicated that a major advantage of solfège versus more 'conventional' aural training/ theory lessons was the fact that it was based on singing: both because it made it more enjoyable and because they considered the ability to sing accurately important for their musicianship. Even students who were very uncomfortable with singing at the start of the lessons, were convinced, by the end of the course, of the vitality of song in performance practice and

had improved dramatically in using it as a medium of musical communication. Students who were enrolled in the Jazz and Popular music programme commented that solfège lessons had improved the level of communication between them and their bandmates; one student also reported that solfège had improved his improvisation skills, as it had helped him 'internalise the sound', so that he could play, to a degree, what he 'heard internally'.

The positive reaction to solfège instruction led to the decision for lessons to be on offer in the next academic year as well: a total of 10 lessons were offered during the first semester, while the duration of the lessons was extended to 90 minutes, so that there was sufficient time for rhythm instruction and harmonic dictation as well. A total of 30 students, divided in two groups, completed a one-semester course and took the exam at the end, which comprised a written and an oral element; feedback for these lessons was, once more, extremely positive.

A general conclusion that can be safely drawn from solfège lessons is that solfège is a natural way to teach theory and aural skills and is appealing to most students. Solfège singing also makes these skills more directly transferable; many students commented on the immediate applicability of solfège singing in their performance practice and some even observed that solfège would be helpful to young children learning an instrument. Although more specialised research is required in order to verify such claims and to determine exactly which aspect of solfège makes it efficient, reports from the present study indicate that the **musical relevance of solfège syllables** (see Chapter 1.2) might be a crucial factor making it extremely practical, easy to learn and enjoyable for music students.

2.2. The pilot study

2.2.0. Aims and hypotheses

The aims of the pilot study were as follows:

- To check the validity of the procedures followed and of the materials used in the experiment.
- To determine whether musicians' scores on working memory span tests can be correlated to their performance scores on music memorisation tasks.
- To determine whether the use of solfège affects the performance of music memorisation tasks.
- To test whether the use of solfège renders memorisation more effective and durable over longer time-spans.
- To find out if the period during which solfège skills are acquired (early childhood or maturity) is crucial to the use of those skills in memorisation tasks.

In order to meet the last three aims, the following hypotheses were tested:

1. Musicians with higher working memory spans will yield higher scores in music memorisation tasks than musicians with lower memory spans.
2. Musicians who use solfège will yield higher scores in music memorisation tasks compared to musicians not using solfège.
3. Musicians who use solfège will yield higher scores in music memorisation tasks after longer time spans compared to musicians not using solfège.
4. Musicians who have learnt solfège at an early age are more likely to use it in music memorisation tasks than musicians who have learnt it during adulthood.
5. Musicians who have learnt solfège at an early age and use it to perform memorisation tasks are more likely to yield higher scores in the tasks than musicians who have learnt solfège during adulthood ad use it to perform the same tasks.

2.2.1. Method

2.2.1.1. Design

The pilot study employed a 3x3x2 mixed design, with three independent, between-subjects factors: i) solfège knowledge, with three levels: solfège, non-solfège and taught solfège, ii) instrument type, also having three levels: singers, string/wind players and pianists; and iii) type of piece tested, with two levels: tonal and atonal piece. The first variable, solfège knowledge, also functioned as an independent within-subject variable for the experimental group. The dependent, within-subject variable was performance score on the music memorisation task. The dependent, between-subjects variables were i) performance scores on the music memorisation task and ii) performance scores on the working memory span tests.

2.2.1.2. Participants

13 students participated in the pilot study, all studying music in the University of Hull. They were divided into three groups:

a) The solfège group, which included 2 international students from countries where solfège is taught and used as part of the music education system. Both students in the groups were females; one was a wind player and the other a singer.

b) The experimental group, which included 6 English and one non-English student who were not previously familiar with solfège. These students followed an intensive solfège course over the spring semester of the academic year 2008–2009 in the University of Hull. There were 3 pianists, 2 wind players, 1 string player and 1 singer in this group.

c) The control group, consisting of 4 English students, unfamiliar with solfège. This group featured 1 pianist, 1 wind player, 1 string player and 1 singer.

All students volunteered to do the experiment and signed consent forms for the music memorisation and automated working memory tests, as well as for participation in the solfège classes.

2.2.1.3. Materials

The memorisation stimuli for the pilot study included three tonal and two atonal pieces, composed specifically for use in this study (see Appendix 2):

1. T1: Tonal piece, 4/4 rhythm, major key.
2. T2: Tonal piece, 4/4 rhythm, minor key.
3. T3: Tonal piece, 6/8 rhythm, major key.
4. A1: Atonal piece, 4/4 rhythm, no key signature.
5. A2: Atonal piece, 4/4 rhythm, no key signature.

Each tonal piece for vocalists, string and wind players had a total of 62 notes, whereas the stimuli for pianists were 108 notes; the key in tonal pieces was different according to the range of the instrument/voice. The two atonal pieces had a total of 41 notes for all vocal and instrumental groups apart from piano players, who had 86 notes in each piece.

All participants were asked to complete two questionnaires, one for the tonal and one for the atonal piece, regarding their music studies, memorisation habits and views towards the task they had just completed (see Appendix 3). The experimental group was also asked to complete an anonymous feedback form, regarding their views on solfège training (see Appendix 1).

All participants were recorded using an Olympus VN-6500 PC digital voice recorder with an additional Olympus ME15 electret condenser microphone. Music memorisation tasks were run in a university practice room in the Department of Drama and Music, equipped with an acoustic upright Yamaha piano (model U3).

The working memory tests used to measure the subjects' operation span and reading span were Automated Operation Span (Aospan) and Automated Reading Span (Arspan) tests respectively (Unsworth et al, 2005), written in E-Prime Version 1.0 and administered to participants in a computer lab in the Department of Psychology, using a regular computer monitor.

2.2.1.4. Procedure

The study involved completion of two tasks, a music memorisation task and a working memory task. The procedure for both tasks is detailed below.

Music memorisation task

An overview of the memorisation task is provided below, followed by a detailed description of all the separate steps:

Phase 1 (February 2009)

1a. Practice and 1st performance

> Learn T1 (10 minutes) – Perform T1 from memory. Complete questionnaire for T1.

1b. (1 hour later) Recall practice and 2nd performance

> Practise T1 without score (1 minute) – Perform T1 from memory.

1c. (24 hours later) Recall practice and 3rd performance

> Practise T1 without score (1 minute) – Perform T1 from memory.

Phase 2 (Experimental group only – February to May 2009)

Intensive solfège programme led by the researcher (see section 2.1.2).

Phase 3 (May 2009)

3a. Practice and 1st performance

> Learn T1 (10 minutes) – Perform T1 from memory. Complete questionnaire for T1.

3b. (1 hour later) Recall practice and 2nd performance

> Practise T1 without score (1 minute) – Perform T1 from memory.

3c. (24 hours later) Recall practice and 3rd performance

> Practise T1 without score (1 minute) – Perform T1 from memory.

Phases 1 and 3 were repeated once more, in February and in May, for each participant, with the atonal piece. Half the participants received the tonal piece first and the other half received the atonal piece first. T1 and T2 are used in the above descriptions as examples; the subjects in the study were given any combination of two different pieces between T1, T2 or T3, for the tonal piece, and A1 or A2, for the atonal piece.

Phase 1a: At the start of the session, the participant was provided with explicit instructions, written and oral, and was encouraged to ask any questions s/he had regarding the procedure. After all queries had been clarified, the participant was left in the practice room alone for 10 minutes, with the instruction to memorise the piece as best as s/he could, using any method s/he wished. A clock was visible to the participants during the 10-minute period for time-management purposes. After 10 minutes the researcher entered the room, took the music away and sat down waiting for the participant to perform the piece from memory – the instructions stated that the participant was not to engage in any form of discussion with the researcher until after completing the memorised performance of the piece. The instructions encouraged subjects to perform the whole excerpt on the first attempt, without restarting if mistakes occurred; if, however, they felt they needed to start from the beginning, they were instructed to do so with the least amount of time possible between attempts. In the case of mistakes and omissions the subjects were instructed to correct only if certain and otherwise proceed with the rest of the piece. The 10-minute practice session and the performance were recorded. After the end of the memorised performance participants were instructed to *not* think about or rehearse the memorised piece, either physically or mentally, until their next session in one hour.

Phase 1b: The participant received written and oral instructions to rehearse what s/he could remember from the piece for 1 minute; performance instructions remained the same as for Phase 1A. Both the 1-minute practice session and the memorised performance were recorded.

Phase 1c: Participants followed the same procedure and instructions as in Phase 1B. After the end of the performance, participants were assured they would not have to perform the piece again.

The procedure and instructions in Phases 3a, 3b and 3c were the same as in Phase 1a, 1b and 1c, respectively.

An obvious discrepancy may be apparent at this point: only two different tonal and atonal pieces are required so that each participant would receive a different piece each time they repeated the experiment in February and in May; yet, for the tonal piece there were three different options. The reason for the existence of a surplus piece was that, prior to the pilot study, a total of 5 tonal and 5 atonal pieces were screened by professional musicians, to whom the aims and the setup of the experiment had been explained. All 5 atonal pieces were judged to be appropriate for the purpose of the experiment, so two pieces were selected at random; the tonal pieces, on the other hand, received varied and, in some occasions, contradictory reviews, which made selection harder. One tonal piece (T5) was considered almost unanimously to be too difficult and was subsequently eliminated; another piece (T1) was unanimously considered appropriate, while the remaining three received mixed reviews. As this was a pilot study, it was considered preferable to put two of the three ambiguous pieces to test, rather than randomly selecting one of them to use alongside T1. Had the number of the pilot study participants been bigger, all four pieces, apart from the extremely difficult one, would have been tested in order to select the most appropriate stimuli for the main study. Given the circumstances, however, the next best solution was to select two of the ambiguous pieces (T2 and T3) and make an informed decision about the main study using the feedback and results from participants in the pilot.

Working memory tasks

The participant came at an appointed time and was given brief verbal instructions regarding the nature of the task, which was completely mouse-driven, with no need for interference by the experimenter. The task itself was broken into three practice sessions, after which the main task took place. For the Aospan test, participants were presented with a series of letters; after the presentation of each letter, they had to perform correctly a very simple mathematical calculation presented on screen (e.g. $(1*2) + 1 = ?$). At the end of the series of letters and mathematical operations, they had to recall the letters presented to them (ranging from 3 to 7), in the correct order, by choosing them from a list of on-screen options. The same procedure was followed in the Arspan test, only this time after the presentation of each letter the participant had to read a short sentence (10–

15 words) and decide whether or not it was nonsensical. Again, after completing a series of such reading tasks, interspersed with presentation of letters, they had to recall all letters presented, in the correct order. For a full account of the Aospan and Rspan tasks see Unsworth, Heitz et al (2005).

For all completed tasks, the program reported five scores to the experimenter: Operation span (Ospan) or Reading span (Rspan) score respectively, total number of correct letters, math or reading errors respectively, speed errors and accuracy errors. The Aospan score has been found to be a reliable indicator of working memory (WM) capacity (Unsworth et al, 2005); WM capacity, in turn, reflects the ability to maintain information in an active state, so that they can be easily retrieved (Engle, 2001). WM span scores in general are used as indicators of higher order cognition (Unsworth et al, 2005), an example of which is the performance of memorised music (Williamon and Valentine, 2002). Music memorisation places a significant load on WM, as information about the serial organisation of the piece needs to be constantly active, while the performer is coordinating a series of complex motor skills to produce the musical material; thus, it seems plausible to hypothesise that higher span scores will function as predictors of accuracy in a music memorisation task with a restricted timeframe.

All participants completed both the Aospan and the Arspan tests, in random order; in the analysis for the present study, only Ospan and Rspan scores were used: these scores represent the sum of all perfectly recalled sets.

2.2.1.5. Transcription and scoring

All six performances of each participant (three memorised performances for the tonal and three for the atonal piece) were transcribed from the voice recorder to manuscript paper. Manuscript paper was preferred to using music notation software for the transcriptions because of the complexity of some of the memorised performances. As a general rule, the more accurate the memorised performance, the simpler the transcription; poor performances normally contained changes in rhythm (bars containing less or more beats than the ones dictated by the time signature), changes in tempo (ritenuti, accelerandi, pauses), more extramusical characteristics (verbal

comments, laughter) and more repeats of the same material within phrases or bars when the participant was unsure of the music being performed. When entered in music notation software, all the above elements result in an overly complex score, which is extremely hard to read and even more difficult to mark; in addition, creating a score which will include details of all changes and mistakes compared to the original is more time-consuming than creating a hand-written score. A sample transcription of a memorised performance is shown in Figure 2.5:

Figure 2.5: Participant no. 9 (cello), T1, 1st performance.

In addition, the 10-minute private practice sessions, two for each participant, were transcribed into text, as well as the two 1-minute recall sessions for each piece.

The scoring method followed for the evaluation of memorised performances is described in detail in section 2.3.

2.2.2. Results

Although all performance transcriptions were evaluated following the scoring system described in 2.3, the limited number of participants in this study made it impossible for any kind of analysis to yield reliable results, which could also be generalised. Especially the small sample of solfège participants was found to be non-representative at all: the singer ranked consistently among the highest scores (above 190 points) in all performances, for both tonal and atonal pieces, when the mean scores for the other 2

singers in the study were 105.98 and 131.93 points for the 1st performance of the tonal and atonal piece respectively. On the other hand, the wind player from the solfège group produced very discrepant performances, ranging from 0 to 196, depending on the piece.

Despite the fact, however, that performance scores could not be used for analysis, the answers from the questionnaires were more informative. All subjects were asked to state the difficulty of the piece they performed each time, in terms of technical difficulty and in terms of memorisation difficulty. Out of the 5 pieces used – 3 tonal, 2 atonal – 2 tonal and both atonal pieces got similar, varied ratings for both their technical and their memorisation difficulty; for the first tonal piece (T1), however, the answers unanimously stated that it was extremely easy, both in terms of technical difficulty and in terms of memorisation difficulty (see Figure 2.6):

Degree of tchnical difficulty

Degree of memorisation difficulty
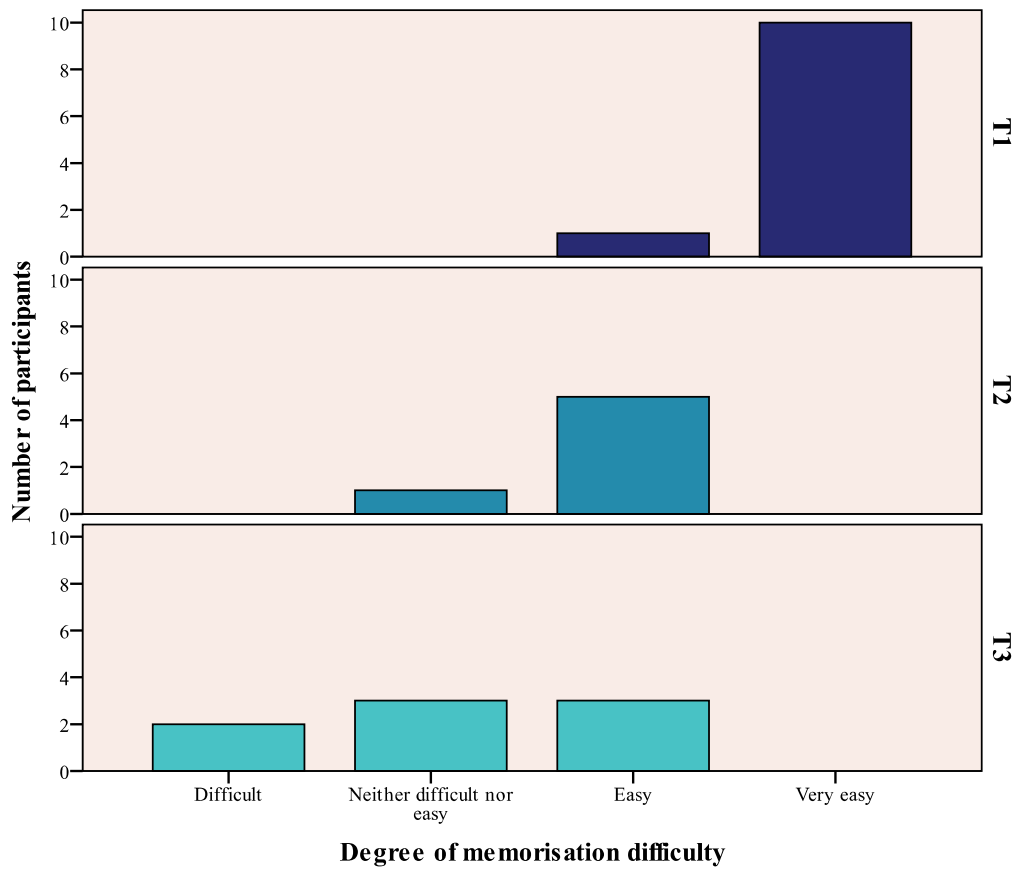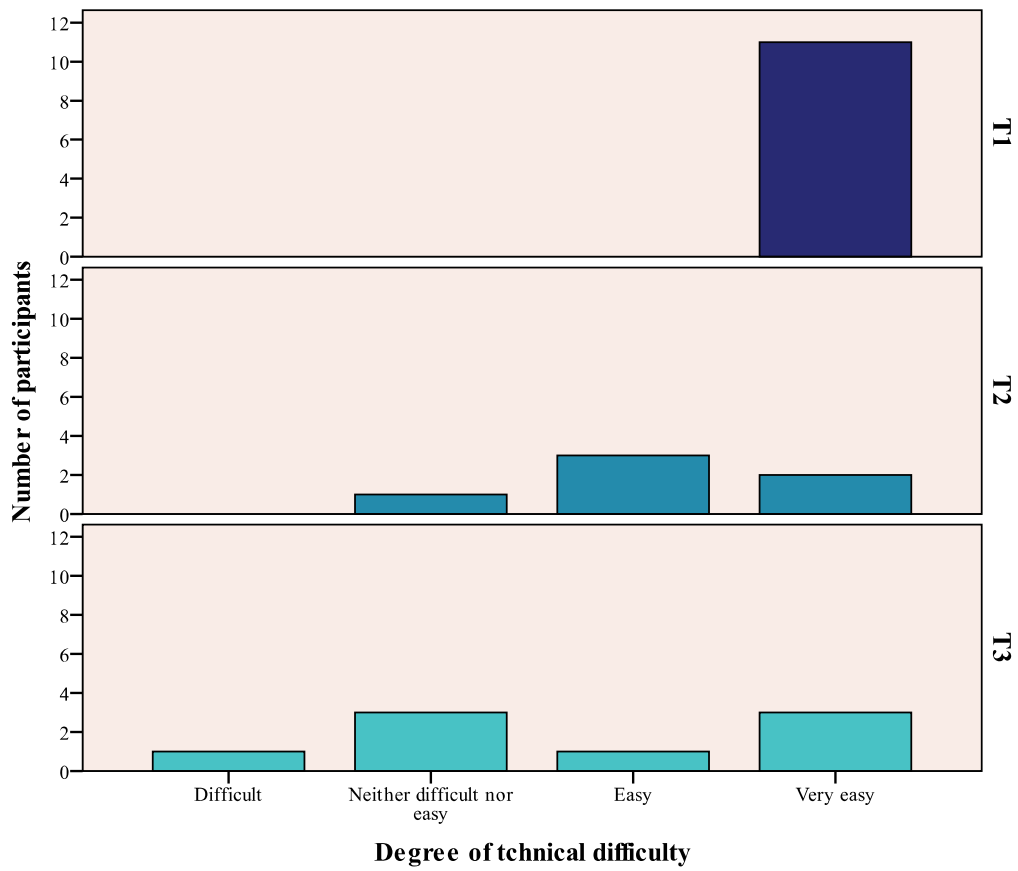
Figure 2.6: Participants' responses regarding the level of technical difficulty and memorisation difficulty for the three tonal pieces. None of the participants selected the 'Very difficult' option, for either technical or the memorisation difficulty of the pieces.

Mean scores for T1, T2 and T3 did not follow the same pattern: as shown in Table 2.3, T2 seems to be the piece that yielded the highest scores. It is important to note, however, that the three participants who consistently produced the highest scores across performances, all happened to get tested on T2, either during Phase 1 or during Phase 2. On the other hand, the two participants with the lowest scores on all performances never got tested on T2; given the small size of the sample, such coincidences are believed to have skewed the results considerably.

Descriptive statistics for the 1st performance of the tonal piece

T1 **(N=9)**	**Mean**	**83.39**
	5% Trimmed mean	**84.12**
	Range	**46.08**
	Skewness	**- .676**
T2 **(N=5)**	**Mean**	**98.36**
	5% Trimmed mean	**98.39**
	Range	**3.91**
	Skewness	**- .663**
T3 **(N=6)**	**Mean**	**81.91**
	5% Trimmed mean	**82.68**
	Range	**47.01**
	Skewness	**- .973**

Table 2.3: Mean scores for the 1st performance of the tonal piece. Note that two participants, who were due to perform T2, failed to attend their experimental session and, consequently, T2 was tested by fewer subjects (N=5).

Subjects from the experimental group were also asked to complete an anonymous feedback form regarding their experience with solfège classes and their general views on solfège as a music education method. Out of 7 students participating in the solfège classes, 3 stated they felt they had enough time to learn solfège, while the other 4 were divided between needing another 2 to 4 weeks and 1 semester to feel they knew solfège satisfactorily. All 7 participants stated that solfège learning improved some aspect of their personal music practice; what they mentioned had improved was, in order of frequency, their aural skills ($N = 7$), their sight-reading ($N = 4$), and their intonation ($N = 1$). Regarding whether or not solfège might be incorporated into a University curriculum, 1 student was neutral while the other 6 were positive towards the idea; on the other hand, on the question about solfège incorporation in the school curriculum, all seven students agreed, with 5 students answering they 'approve strongly' and 2 they 'approve' of such an idea.

Out of the 7 students participating in the solfège classes, only 4 were asked to use it in the second round of memorisation tasks in May; the students were screened according to their results in a solfège exam which had taken place a week before the second round of the experiment. Four students who scored over 50% in the solfège exam were instructed to use it as much as possible in the memorisation tasks; the rest did not receive particular instructions and did not use solfège. The transcription of the practice sessions, however, revealed that, out of the 4 participants instructed to use solfège, only one had actually used it, at least in a way that was audible from the recording. Despite this finding, all 4 participants completed the questionnaire as if they had used solfège and, in the question 'How did solfège affect your memorisation?' regarding the tonal piece, 2 participants found a positive effect, one found a negative effect and one found no effect at all. It is worth mentioning that the participant reporting the negative effect was a singer and also the one who had actually used solfège audibly during the practice session; the negative effect was explained in a subsequent question, where the subject stated that he had found solfège helpful but too time-consuming and ultimately detrimental, given the restricted timeframe. The exact same responses were recorded in the corresponding question regarding the atonal piece; this time, the male singer who reported a negative effect added that he would have liked to use more solfège if he had more time to complete the task.

According to the first hypothesis:

    1. Musicians with higher working memory spans will yield higher scores in music memorisation tasks than musicians with lower memory spans.

In order to test this hypothesis, the performance scores from Phase 1 of the pilot study needed to be compared with Aospan and Arspan scores. Several problems arose in regard to these comparisons:

1) The Aospan and Arspan tests used were in English and therefore appropriate to use only with English participants. This meant that both solfège participants, as well as one participant from the experimental group, were automatically excluded from the automated working memory tests, lowering the number of participants to 10.

2) Questionnaire responses revealed that participants had unanimously found T1 easier than T2 and T3; the fact that this had not been validated from performance scores, whereas independent judges had considered T2 and T3 harder than T1, led to the conclusion that the overall consistency and reliability of the performance scores had been compromised.

For the reasons described above, instead of attempting to compare covariances between the WM tests and the performance scores, the WM test results were first plotted in a line graph (Figure 2.7) against the memorisation scores for the 1st performance of the tonal and the 1st performance of the atonal piece. As the maximum possible score was 100 for the music tests (see section 2.3) and 75 for both WM tests, all Aospan and Arspan scores were also converted to percentage scores for this comparison:
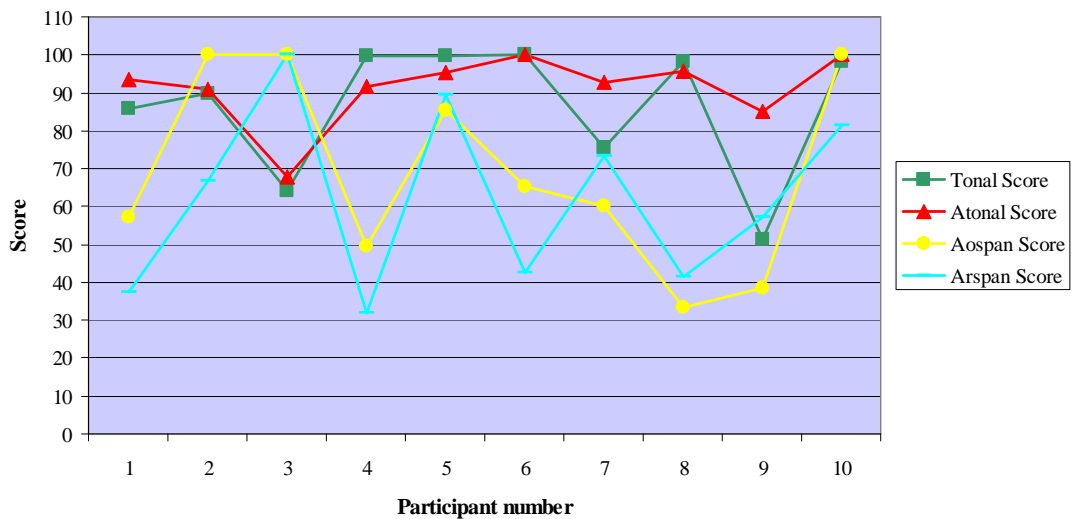
Figure 2.7: The triangle, square, circle and dash shapes demonstrate the value of 1st performance score for the atonal piece, the 1st performance score for the tonal piece, the Aospan and the Arspan score respectively, for each participant.

The lines in Figure 2.7 suggest that the working memory scores and the performance scores did not correlate in a consistent manner, although, in most cases, they seem to be similar within groups: participants generally produced similar results in the memorisation task for the tonal and the atonal piece, and similar results in the two working memory tests. The most notable cases are summarised below:

- Participant no. 3 scored maximum points in both working memory tests. His performance score, however, was the lowest for the atonal piece and second lowest for the tonal piece.

- Participant no. 8 scored maximum points in both the tonal and the atonal memorisation tasks, whereas her Aospan score was the lowest and her Arspan third lowest.

- Participant no. 10 scored maximum points in all but the Arspan test.

- Participants 4, 5 and 6 scored above 90% for both pieces, while their Aospan scores varied as much as 44% and their Arspan scores as much as 57%.

According to the second, third and fifth hypothesis:

2. Musicians who use solfège will yield higher scores in music memorisation tasks compared to musicians not using solfège.

3. Musicians who use solfège will yield higher scores in music memorisation tasks after longer time spans compared to musicians not using solfège.

5. Musicians who have learnt solfège at an early age and use it to perform memorisation tasks are more likely to yield higher scores in the tasks than musicians who have learnt solfège during adulthood ad use it to perform the same tasks.

It was not possible to conduct reliable statistical tests for these hypotheses, due to the extremely limited number of solfège participants and the apparent inequality in the difficulty level of the memorisation stimuli. It is worth noting that the solfège singer scored a minimum 44.21% higher than the mean of the two non-solfège singers in the performances of the tonal piece and a minimum 30.53% higher than the non-solfège singers in the performances of the atonal piece; the wind player form the solfège group yielded extremely discrepant scores, varying from 0 to 196, so that no comparisons could be made to the instrumentalists in the non-solfège group. Despite these problems, however, pilot performance results led to certain very interesting conclusions and corrections to be made in the design of the main study; these are presented in the Discussion.

According to the fourth hypothesis:

4. Musicians who have learnt solfège at an early age are more likely to use it in music memorisation tasks than musicians who have learnt it during adulthood.

Out of the seven students participating in the solfège training programme, only four managed to score above 50% in the solfège exam; moreover, the recordings of practice sessions of those who scored over 50% in the exam revealed that they had not been able to incorporate solfège effectively in their memorisation strategy. These findings provide tentative support for hypothesis no. 4, however, a much bigger sample is necessary in order to be able to generalise any conclusions drawn.

2.2.3. Discussion

One of the aims of the pilot study was to check the validity of the procedures followed and of the materials used in the experiment. This aim was met, as, according to the results of the pilot, several changes were made in the design of the main study. The main points emerging from the pilot are summarised below:

2.2.3.1. Review of Participants

Bearing in mind sample size constraints, analysis showed than the categorisation of participants in: singers, string players, wind players and keyboard instrument players, was meaningful; it also revealed the additional possibility of the further regrouping of string and wind players in one group, being instrumentalists who normally have to perform a single melodic line. The latter grouping scheme was adopted when recruiting participants for the main study: target subject numbers in the vocal and the string/ wind group were approximately equal, while the target number of keyboard instrument players was slightly smaller, due to the nature of the hypotheses – solfège was expected to be least appropriate for use and least effective for pianists than for the other two groups, owing to the polyphonic element in the musical material.

A tuned percussion student who had volunteered to participate in the pilot study had to be rejected due to practical difficulties with arranging experimental sessions in the percussion practice room and moving the instruments; to avoid such complications, percussionists were excluded from the main study as well.

2.2.3.2. Review of Materials

Musical stimuli
The original aims regarding the pieces used in the pilot were: a) to confirm that the pieces deemed appropriate by independent judges (A1, A2 and T1) were appropriate for the main study and b) to select which one out of T2 and T3 would be most appropriate

to use alongside T1. During the pilot, however, it was made apparent that the experimental process ran more smoothly with the tonal pieces than with the atonal, owing to the one extra tonal piece used. This happened because, even though there were only a few participants, they tended to make back-to-back appointments for their experimental sessions: as a consequence, it was often the case that one participant would be waiting for their turn outside the practice room, while another was inside the room, practising or performing a piece; although practice rooms were relatively sound-proof, it was relatively easy for people outside practice rooms to overhear the music produced inside. Having three options for the tonal pieces made it much easier to avoid situations in which the participant waiting outside would be tested on the same piece as the one performed by the participant inside the practice room; in some occasions, this was not possible for the atonal pieces. Taking into consideration that the number of participants in the main study would be considerably bigger and that, for the experiments run outside the UK, there would be an even stricter timeframe within which all appointments should take place, it was decided best to have a selection of 3 tonal and 3 atonal pieces to use in the main study.

Regarding the selection of the pieces, the first tonal piece (T1), which had been the only one deemed appropriate by independent judges, had been found too easy and was therefore substituted by one of the other tonal pieces which had received mixed reviews from the independent judges. The new piece was also in major key and 4/4 rhythm, but did not contain extensive scale passages, which was generally considered to be the main feature of T1 that made it easier than the other two.

The two atonal pieces were kept the same for the main study, apart from the pianists' stimuli in the left hand for A2. This was also proved excessively easy, especially for pianists employing the strategy of learning the right hand separately from the left, as it contained sequences which, by themselves, sounded as having a tonal centre. The result was that some pianists concentrated on learning the – easier – left hand part and remembered it in isolation, without being able to remember more than a bar from the right hand. Thus, the left hand part in A2 was corrected for the main study. Taking into account that the original 5 atonal pieces had received similar reviews by independent judges regarding their overall difficulty and appropriateness for the task at hand, one of

them was added in the material for the main study, in order to have a total of 3 atonal pieces.

Working memory span tests

The Aospan and Arspan test scores, measuring the subjects' operation span and reading span respectively, have been proved to be reliable indicators of a subject's working memory span, which in turn has been shown to predict performance in higher order cognitive tasks (Conway et al 2005, Unsworth et al 2005). These test scores, however, could not be correlated to music memorisation task performance scores in any way. That could happen for any of the following reasons:

1. Music memorisation does not fall within the domain of higher order cognitive tasks.
2. The working memory tasks used were inappropriate for the task at hand.
3. The music memorisation performance scores were inaccurate.

As discussed in Chapter 1.3.3, the first proposition is most definitely false; music memorisation has been generally admitted to be a highly refined cognitive task, demanding the fine tuning of both cognitive and sensory-motor skills. The remaining two, however, are plausible; participant feedback as well as the dataset itself reveals that there are certain theoretical and methodological problems in the pilot design.

The problems of the scoring system used in performance evaluation are extensively discussed in the subsequent section (Chapter 2.3); the pilot performances, however, could have been re-scored according to the improved evaluation system described in section 2.3.3. The reason this was not done was because the main problem was believed to be a significant difference in the level of focus between the working memory tests and the music tests. The music tests from which the memorisation scores resulted were targeted to skilled musicians, evaluating an extremely domain-specific set of skills, relying on deep structures; on the other hand the Aospan and Arspan tests tested a very broad, surface range of skills and can be registered to anyone. An illuminating example of the resulting imbalance is that the participant who had mentioned s/he was very competent in mathematics and was very confident about taking the Aospan test scored, indeed, the highest in the test; on the other hand, the participant who was least confident and had previously reported that she could not read efficiently from a computer screen

had one of the lowest scores in the Arspan test. Another flaw in the use of the particular automated tests as indicators of musical working memory span is that, there being a certain limit in the maximum points a subject can score, there was a number of participants that achieved a maximum score in both tests; this, of course, does not mean that the particular subjects had the exact same level of skill. In both the above examples the participants' performance in the Aospan and Arspan tests seems to be related to a broad set of cognitive skills, therefore the correlation of these results with performance in music tests constitutes a gross generalisation and over-simplification. Although these particular automated tests are reliable when used to categorise subjects in high-span and low-span groups, as is one of their most common uses, it was evident that a different, domain-specific test needed to be designed to measure music working memory span. Such a test was designed for the purposes of the main study, using the same platform and general format, but using music-specific stimuli, so that the scores could be compared to music memorisation scores.

2.2.3.3. Review of Procedures

The main changes resulting from the pilot were the following:

- **Arrangement of first memorisation session and subsequent recall session.** The arrangement of the first two sessions for each piece happening an hour apart created numerous problems, both practical and methodological. In practical terms, it proved extremely difficult to find dates at which the participants were free at the exact times specified in the experiment design (30 minutes, 5 minutes after one hour and 5 minutes after a day). In addition, in several cases it proved impossible to keep participants from involuntarily rehearsing the piece mentally during the hour between the first and the second session. For these reasons, the second session was incorporated in the first for the main study; participants were kept in the practice room for approximately 40 minutes, during which the 10-minute memorisation session, a 15-minute distraction period and the recall session and performance took place. A similar design has been also followed in Ginsborg and Sloboda (2007), where a 10-minute interview was carried out between sessions; this design ensures that participants are kept from rehearsing the piece, while at the same time long-term memory is being tested after the 10

or 15-minute period has passed. This design offered the chance to conduct the first two performances of the piece as smoothly as possible; the potential problems, however, arising from the 24-hour break between the 2nd and the last recall of the piece were more difficult to resolve. Since it was impossible to control participants' actions and thoughts between the first and the last session, an attempt was made to preserve environmental factors of the experimental procedure, in order to avoid changes in performance resulting from the change of room or instrument, in the case of pianists (Mishra and Backlin, 2007). The restrictions posed by the experimental design between the first and last sessions were taken into consideration in data analysis for the main study.

- **Duration of the recall practice sessions.** Recordings showed that, in most instances, the 1-minute interval the participants had in order to refresh their memory of the piece was not enough. Apart from participants who were absolutely confident and performed the piece with over 90% accuracy, all other subjects used up the 1-minute period either frantically trying out bars which they could not remember, or playing almost nothing at all, trying to concentrate. Durations of 1.5, 2, 2.5 and 3 minutes were piloted to a small sample and the 2-minute recall practice duration was kept in the main study.

- **General instructions for performance.** Many of the subjects who had low scores had yielded performances which sounded more like practice sessions and, in some cases, were as long as 3 minutes, when the average duration for the particular 8-bar pieces performed from the beginning to the end at slow tempo was 30 seconds. In order to avoid such phenomena in the main study, written and oral instructions were added, in which subjects were instructed to view the occasion as a real performance for an audience and thus react to any mistakes in the same manner they would if performing the piece in public.

- **Scoring system.** Evaluation of pilot performances proved exceptionally complex and it thus underwent several changes and stages. A full account of the scoring of pilot performances is given in the following section (2.3).

It is also worth noting that the audio recording of participants was decided to be kept in the main study, instead of being switched to video recording, as it provided a balance between the quality of recorded data and extent of interference with participants' performances: according to their responses in the questionnaire, the number of

participants that were significantly distracted by the fact they were being recorded was 2 for the tonal and 3 for the atonal pieces, out of a total of 13 participants. These numbers suggested that, however small, there was still a portion of students whose performance results might have been skewed due to anxiety caused by recording; informal feedback from all subjects in the pilot indicated that video recording would have been a more severe distraction for even more participants. The quantity and quality of the data extracted from audio recordings, on the other hand, was, in the vast majority of the cases, sufficient for an effective analysis of results; although in some instances, especially in the cases of extended silence intervals, it was not possible to decipher with certainty what the participant was doing, these cases were extremely few and also relatively easy interpreted using contextual information.

2.2.3.4. Review of Solfège training

Perhaps the most important point emerging from the pilot was verifying and understanding the process of teaching solfège to adult musicians. Because solfège, and especially fixed-do solfège, is normally taught to children who are just beginning to learn music, one of the main questions was how much time it would take adults to adjust to this new way of, initially, reading music. Solfège lessons, practice session recordings and questionnaire responses revealed that, as in all fields of learning, the duration depends quite a lot on the individual; out of seven students only four managed to score above 50% in a solfège exam and even those did not seem able to employ solfège productively and incorporate it in their memorisation strategy. This seems to converge with hypotheses number 4 and 5:

4. When solfège is learnt at an early age it is more likely to be put in use in music memorisation tasks.

5. When solfège is learnt at an early age it is more likely to be effective when used in music memorisation tasks.

A larger sample, however, both for the taught solfège and for the solfège groups, is definitely required in order to be able to reach any definitive conclusions.

In summary, the first hypothesis was not supported from the data; due to methodological limitations, however, including inadequate number of participants and the lack of fit of the WM tests used to the purpose of the study, no definitive conclusions were drawn regarding the link between WM span and performance on music memorisation tasks. Hypotheses 2 and 3 could not be tested due to lack of subjects using solfège; finally, transcription evidence suggested that Hypotheses 4 and 5 could be tentatively considered valid, although a larger sample would be required in order to draw safe conclusions.

2.2.4. Conclusions from the pilot

Although most hypotheses could not be efficiently tested, mainly due to the limited number of participants but also due to several methodological issues that emerged, the pilot study achieved its first and most important aim: it verified the validity of certain procedures followed in the experiment and produced feedback which changed certain other procedures as well as materials used in the main study.

Moreover, the complexity of the design and several problems that arose during the pilot led to the decision of divide the main study in three separate experiments:

- Experiment 1 employed a between-groups design, in order to test the differences in memorised performance scores produced by solfège and non-solfège subjects.
- Experiment 2 employed a repeated-measures design, which compared performance scores yielded by non-solfège subjects before and after they received solfège training.
- Experiment 3 included three sets of short experiments, all following a between-groups design: each one was designed to test the function of solfège syllables and letter names in working memory by comparing memorisation task performance in solfège and non-solfège subjects, respectively.

2.3. Scoring the memorised performances

2.3.1. Preliminary marking

All six memorised performances of each participant – three for the tonal and three for the atonal piece – were marked according to pitch accuracy, rhythm accuracy, structural accuracy and fluency; the overall score was the sum of the four scores. The marking system used to evaluate pitch and rhythm accuracy was based on the verified scoring algorithm provided by Gilman et al (2002) for assessing sight-reading performances, with some modifications that were necessary in order to accommodate the lack of a set metronome beat in the memorisation tasks of this study, as well as the inherent differences between sight-reading and performance from memory; Segalowitz et al (2001) have underlined the difficulties presented in the quantitative scoring of fragmentary performances. Following Gilman et al's model, each note played received a maximum of 1 point for pitch accuracy and 1 point for rhythm accuracy. In case of mistakes in pitch, 1/12 of a point was deducted for each semitone difference between the note played and the original; in case of rhythm errors, 1/4 of a point was deducted for each semiquaver difference between the note played and the original note. Missing notes received zero points and in the case of extra notes within the meter (e.g. two quavers instead of a crotchet) the penalty was equal to the penalty for the rhythmic difference between the original note and the first of the extra notes (e.g. - 0.5 point in the case of two quavers in the place of a crotchet). In the case of extra notes played to fill in a memory gap, the notes were counted as missing (zero points received) without further penalties regarding rhythmic values. Both for pitch and for rhythm accuracy, the minimum score for a note could be zero; no note could receive a negative score.

The structural accuracy score was initially designed to measure the percentage of the bars remembered and played correctly and, as such, was calculated per total number of bars in the piece. Each correct bar in the correct position received 2 points, so the maximum score for each piece was 16 points. If a correct bar was played in a wrong position one point was deducted and if a wrong bar, in terms of harmonic structure or contour, was played in order to substitute a forgotten bar, again one point was deducted.

If more than 8 bars were played, 0.5 points were deducted for each extra bar. In the case of missing phrases or bars, 0.5 points were deducted for each crotchet missing.

The fluency score resulted from the need to distinguish between performances that were equally accurate in pitch, rhythm and structure but had many differences in the ease with which they had been delivered and, thus, was first designed as a measure of flow. Fluency was calculated per total number of beats played, the beat in 4/4 metre being the crotchet. Each beat played received one point and points were deducted according to the following table:

Type of error	Penalty
Apoggiatura	- 0.25
Hesitation (momentary)	- 0.25
Instant correction of wrong note	- 0.5
Delay (between 0.5 – 1 beat)	- 0.5
Delay (more than 1 beat)	- 1
Repeat of a note/motif	- 0.75 per note/motif
Restart a phrase/bar	- 0.75
Ritenuto/accelerando	- 0.25 per beat

This version of the marking system did not penalise out of tune notes, following the premise that tuning issues were irrelevant to memorisation competence. All four scores (Pitch, Rhythm, Structure, Fluency) were added up and the result was converted to a percentage score; thus, the maximum score a participant could receive for a single performance was 100 points.

2.3.2. Validation

After marking all performances using the above system, the highest, middle and lowest ten scores for the tonal performances were selected and the respective performances were entered in a CD in random order. The CD was given to two independent judges, who were instructed to rank the performances from highest to lowest according to

memorisation accuracy, disregarding any technical and tuning issues. The two judges were also given the original musical scores of the pieces performed, as well as a separate CD in which the pieces were performed correctly on a piano.

Both independent judges were professional musicians: the first was a female employed as a cellist in a symphony orchestra and as a teacher in a conservatoire, with substantial experience as a soloist; the second was a male, also employed in a symphony orchestra and participating in numerous ensembles, performing music of various genres. Their rankings were compared between them and with the ranking resulting from the scoring system, resulting in the following conclusions:

1. Both independent judges had effectively divided the 30 performances into 3 groups, which coincided with the highest, middle and lowest group resulting from the scores' ranking.

2. The only places where the two judges' rankings were identical were: 1st place (most accurate performance), 5th place, 10th place (boundary of the high-score group) and 30th place (least accurate performance). The score rankings also coincided with the judges' decision for 1st, 10th and 30th place, but not for 5th place: the performance which the two judges had ranked 5th was ranked 6th in the scoring system.

3. Feedback from the judges regarding the task was that it was quite hard and, at times, impossible for them to completely disregard technical and tuning issues. They both stated that the effect of such issues and general fluency issues was particularly strong in poorest performances, while in the most accurate performances it had been easier to overlook flaws in delivery. Their musical intuition dictated that such issues could not and should not be overlooked completely when assessing memorised performance.

4. They both felt that, apart from the 1st, 10th and 30th place, the rest of the rankings were negotiable and stated they would be more comfortable discussing their rankings with peers, as is the norm in exams and competitions. They both stated that, in many cases, even after having repeatedly listened to some of the performances, they felt that the differences between them were not significant or important, especially since they were asked not to take into account general 'musicality' issues.

5. The biggest variation in the 3 rankings was found in the lower half of the scores, with the exception of the last place. As one of the judges put it: 'When we're bordering the realm of aleatoric music, it really doesn't matter any more. They are all just wrong.'

6. In a general discussion about scoring systems, both judges suggested that they found it hard, if not impossible, to determine a single scoring system that can predict and include every option. They pointed out that they expected a system of marking memorised performances to have some elements in common with systems of marking dictation tests, especially when faced with large-scale errors such as transpositions of passages or metric errors.

Taking these comments into account, as well as the independent judges' rankings, the marking system was revised and enhanced so as to yield results which would be closer to natural musical intuition and judgement. The resulting marking system does not purport to be an all-encompassing matrix for marking memorised performance and the resulting scores may always be subject to criticism, as is very often the case with the evaluation of music performance; it is, however, a system which can be used across a set of performances in order to produce consistent results and, as such, is suitable to serve the purposes of the present study.

2.3.3. Final scoring system

The final system for evaluating memorised performances comprised the same four components as the initial marking system – pitch accuracy, rhythm accuracy, structural accuracy and fluency – with the following changes (for the complete set of specific rules for each score see Appendix 4):

2.3.3.1. Pitch accuracy

The main change in the calculation of the pitch accuracy score (hereon Pitch Score, or PS) was the treatment of 'transposition errors', which led to the addition of the following rules:

- For single wrong notes with an octave difference from the original the penalty equals - 2/12 points.
- For transposed passages the penalty follows the formula:

$$penalty = E_1 + \frac{E_1 \times (N-1)}{2}$$

  where E_1 is the penalty for the first note (1/12 points deducted for each semitone difference) and N is the total number of notes in the transposed passage.

- In the case of erroneous repeated motifs, the most prominent examples found in the left hand of the keyboard stimuli, the penalty is calculated only for the first motif.
- In the case of erroneous doubled notes, a penalty is calculated only for the first note.

2.3.3.2. Rhythm accuracy

The main changes in the calculation of the rhythm accuracy score (hereon Rhythm Score, or RS) were the following:

- Extra notes do not incur a penalty in the rhythm score, but rather in the Fluency Score and, in the case of whole extra phrases performed, in the Structure Score.
- Performing a motif using the correct notes but switching the positions of rhythmic values (e.g. θ θ η instead of θ η θ) incurs half the normal penalty.

2.3.3.3. Structural accuracy

The main changes in the calculation of structural accuracy score (hereon Structure Score, or SS) were the following:

- In the case of extra notes performed as trials the penalty equals half the normal penalty for extra beats.
- Performance of extra bars incurs a penalty of - 0.25 per extra bar.
- Performance of a wrong bar to fill in a memory gap incurs a penalty of - 1.25 points for tonal pieces, while the penalty of - 1 points for atonal pieces is retained.

2.3.3.4. Fluency

The main changes in the fluency score (FS) were:

- Each tempo change incurs a penalty of - 2 points.
- Performance of extra notes or motifs, which are outside the metre, incurs a penalty of - 0.75 points for each note/motif.
- Out of tune notes incur a penalty of - 0.25 each. In case of general tuning issues in a performance, 1 point is deducted from the overall Fluency Score.

After calculating the four sub-scores, PS and RS were added up to create the Technical Score (TS). The three resulting scores, Technical, Structural and Fluency were then converted to percentage scores and added up according to the following formulae:

For TS $\geq$ 72.87 %

$$\text{Overall Score} = \text{TS} + (8.95 \times \text{SS}) + (0.05 \times \text{FS})$$

For TS $<$ 72.87 %

$$\text{Overall Score} = (2.25 \times \text{TS}) + (6.75 \times \text{SS}) + \text{FS}$$

The threshold of 72.87% for the TS was calculated using a combination of individual scores from the pilot and hypothetical scores: certain scores were used as 'landmarks', outlining lower and upper thresholds for low, medium and high-performance groups. Boundaries between low, medium and high-performance groups were selected by the researcher according to the independent judges' markings and reports, as well as based on the researcher's own intuitive musical judgement. Factors in the individual algorithms were determined by a trial-and-error procedure; the resulting Overall Scores led to a ranking of participants which is consistent with musical intuition in the more clear-cut cases of excellent or poor performance and, at the same time, provides a reasonable scoring pattern in the more ambiguous cases. The values of factors in the formulae were selected so that the maximum Overall Score a participant could obtain was 1,000.

Chapter 3

Experiment 1

The effect of solfège instruction on music memorisation: A cross-cultural study

3.0. Aims and hypotheses

Experiment 1 set out to examine the differences in the performance of music memorisation tasks by solfège and non-solfège musicians. The general aims of the experiment were:

1. To examine the effects of solfège use on a music memorisation task performed by adult musicians.

2. To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians.

In order to achieve these aims, the following hypotheses were tested:

1. **Musicians who know and use solfège will perform better in a music memorisation task than musicians who do not know and do not use solfège.**
   The direction of this hypothesis was dictated by the theoretical arguments postulated in Chapters 1.2 and 1.3; solfège was expected to provide a comparative advantage, owing to its dual verbal and musical nature.

2. **The positive effect of solfège utilisation in the memorisation process will be highest amongst singers and string players and lowest in piano players.**
   As solfège is inextricably linked to singing, the presence of polyphony and the inability to sing more than one melodic line simultaneously was expected to eliminate the advantage offered by solfège in pianists' memorisation. The link between solfège and singing was also expected to create problems in its use by wind players: as they are required to use their mouth when performing on their instruments, the practice of solfège was expected to create a distraction, given the restricted timeframe of the task.

3. **The positive effect of solfège utilisation will be stronger for atonal music than for tonal music.**
   Musicians are based on learnt patterns and previously formed expectations to form the memory trace of a piece of music, therefore tonal stimuli were expected to be treated in similar ways by all subjects. Atonal stimuli, on the other hand, were expected to pose more challenges regarding pattern extraction and

grouping; a positive effect of solfège was therefore expected to appear, due to the categorisation possibilities solfège syllables provide (Chapter 1.2).

It is worth noting at this point that, although the above hypotheses were the ones tested for the purpose of the present study on solfège, the quantity and quality of the data gathered during Experiment 1 were offered for much more extensive analysis of numerous variables affecting performance. Although it is beyond the scope of the thesis to provide an exhaustive analysis of the parameters factoring in memorisation performance, the most notable findings will be reported and discussed during the course of the analysis.

3.1. Method

3.1.1. Design

The study employed a between-groups design, with three independent variables:
1. **Solfège use**, with two levels: solfège and non-solfège.
2. **Instrument performed**, with four levels: voice, string, wind, piano.
3. **Type of piece performed**, with two levels: tonal and atonal.

The above independent variables were the ones explored in the process of testing the directional hypotheses stated in the previous paragraph; during the course of data analysis, the possible effect of other independent variables was explored, such as absolute pitch possession, memorisation frequency and memorisation strategy employed in the task. The dependent variable was the final score on the memorised task. All participants were asked to perform the task twice, one with a tonal and one with an atonal piece; the order in which they received the pieces was quasi-random.

3.1.2. Participants

A total of 93 university music students participated in the study; 48 students came from countries in which solfège is taught and used as part of the music education system from the beginning to tertiary education and 45 students came from countries in which solfège is not normally used in music education. The 48 solfège participants included 19 singers, 8 string players, 7 wind players and 14 pianists, whereas in the non-solfège group there were 13 singers, 9 string players, 7 wind players and 16 pianists. All instrumental performers had had over 8 years of lessons on their respective instruments at the time of their participation, whereas singers were required to have either a minimum of 1 year of vocal lessons or a minimum of 5 years of singing in a choir that had classical repertoire and used sheet music. The age of participants ranged from 17 to 57, the vast majority, however, (84.1% of the subjects) were below 23. All participants volunteered for the experiment and signed the relevant consent forms; in addition, 30 students from each group were offered course credit for their participation.

3.1.3. Materials

Six 8-bar pieces, three tonal (T1, T2, T3) and three atonal (A1, A2, A3), created specifically for the experiment, were used as stimuli:

1.   T1 (Tonal piece): 4/4 rhythm, major key.
2.   T2 (Tonal piece): 4/4 rhythm, minor key.
3.   T3 (Tonal piece): 6/8 rhythm, major key.
4.   A1 (Atonal piece): 4/4 rhythm, no key signature.
5.   A2 (Atonal piece): 4/4 rhythm, no key signature.
6.   A3 (Atonal piece): 4/4 rhythm, no key signature.

Each tonal piece for vocalists, string and wind players had a total of 62 notes, whereas the stimuli for keyboard instruments were 108 notes. Each atonal piece had a total of 41 notes for all vocal and instrumental groups apart from keyboard instrument players, who had 86 notes in each piece. The key in tonal pieces was different according to the range of the instrument/voice; atonal pieces for different instruments/voices were also

transposed in order to accommodate different ranges, although this was not accompanied by a key signature change.

The experiments took place in the practice room facilities of the institutions the participants were attending at the time. Apart from reasonable and expected differences between practice rooms in different universities, all rooms had a piano or clavinova; in most cases each participant was able to complete the different sessions of the experiment in the same practice room, but in several instances that was not possible, due to departmental timetabling. All participants were recorded using an Olympus VN-6500 PC digital voice recorder with an additional Olympus ME15 electret condenser microphone.

3.1.4. Procedure

An overview of the experimental design is provided below, followed by a detailed description of all the separate steps:

1a. Practice and 1st performance

   Learn T1 (10 minutes) – Perform T1 from memory.

1b. Distraction (immediately after 1st performance)

   Complete questionnaire for T1 and perform distraction tasks (15 minutes).

1c. Recall practice and 2nd performance (immediately after Distraction)

   Practise T1 without score (2 minutes) – Perform T1 from memory.

2. Recall practice and 3rd performance (24 hours later)

   Practise T1 without score (2 minutes) – Perform T1 from memory.

1a, 1b and 1c happened consecutively, in a single session lasting approximately 45 minutes. Session 2 took place exactly 24 hours after session 1a. All four sessions (1a,b,c and 2) were repeated for the Atonal piece. T1 is used in the above description as an

example; the subjects in the study were given any combination of two different pieces between T1, T2 or T3 and A1, A2 or A3.

1a. Practice and 1st performance (15 minutes)

i) Introduction and instructions

The participant was asked to come in the practice room facilities of the institution they were attending at a pre-appointed time. They were provided written, explicit instructions in their native language and they were encouraged to ask any questions they had regarding the procedure. Communication took place in the participants' native language in all but six cases, in which English was used, spoken fluently by both parties. Solfège participants were specifically instructed to use solfège, for at least part of the piece, in order to memorise it, even if they were not normally used to employing this method.

ii) Practice

The participant was given the piece and was left in the practice room alone for 10 minutes, with the instruction to memorise it as best as s/he could, using any method s/he wished. Solfège participants were encouraged to use solfège for at least part of the piece but, if they did not feel comfortable with it, they were instructed to employ solfège for up to 2 minutes in total. A clock was visible to the participants during the 10-minute period, for time-management purposes, and the 10-minute practice session was recorded.

iii) 1st performance

After 10 minutes the researcher entered the room, took the music away and sat down waiting for the participant to perform the piece from memory. The instructions stated that the participant was not to engage in any form of discussion with the researcher until after completing the memorised performance of the piece. The instructions for the performance itself stated that the participant should regard the occasion as a performance rather than a practice session, so that their actions and reactions to any mistakes should be equivalent to those when performing for an audience. The subjects were encouraged to perform the whole excerpt on the first attempt, without restarting if mistakes occurred; if, however, they felt they needed to

start from the beginning, they were instructed to do so with the least amount of time possible between attempts. In the case of mistakes and omissions the subjects were instructed to correct only if certain and otherwise proceed with the rest of the piece. The performance was recorded.

At the end of the performance, subjects received feedback regarding their response to the task. This was deemed necessary, especially in cases when the subject had produced a poor performance; all subjects were assured that the desirable end product of the task was not an impeccable memorised result, but rather a consistent, continuous performance, regardless of memorisation quality. Moreover, in order to ensure that participants would not feel in any way inadequate or incompetent as a result of their participation in the task, the researcher explained clearly that the aim of the study was mainly to record the strategies and techniques used for memorisation, rather than to simply evaluate the memorised performance.

1b. Distraction task (15 minutes)

After the end of the memorised performance, the participant took a 15-minute 'break' from the task, during which s/he was instructed to avoid rehearsing the piece mentally: s/he was asked to complete a questionnaire and then try to relax, not thinking about the music. The purpose of the questionnaire was to gather information about the subjects' musical background, their memorisation habits and their views on the task they had just completed, while at the same time they were being distracted from thinking about the music. In the time remaining after the completion of the questionnaire, they could choose to solve short puzzles, provided by the researcher, eat sweets, or engage in conversation. The 15-minute period was not recorded.

1c. Recall practice and 2nd performance (3-5 minutes)

i) Recall practice
At the end of 15 minutes the experimenter left the room and the subject was given 2 minutes to rehearse what they could remember from the piece, either physically or mentally, or both, without being given the music again. The 2-minute practice session was recorded.

ii) 2nd performance

After 2 minutes the researcher re-entered the room and sat down waiting for the participant to perform the piece from memory. The instructions were the same as for the 1st memorised performance: the participant was not to engage in any form of discussion with the researcher until after completing the performance, which should not be viewed as a practice session regarding reaction to mistakes and general comportment. The performance was recorded.

2. Recall practice and 3rd performance (3 minutes)

The second session was arranged to take place exactly 24 hours after the first. During the second session, the participant was left alone for 2 minutes in the practice room, instructed to rehearse what they could remember from the piece, either physically or mentally, or both, without being given the music again. At the end of 2 minutes, the researcher entered the room and the subject performed the piece from memory. Both the 2-minute practice session and the performance were recorded.

3.1.5. Transcription and scoring

All six performances of each participant (three memorised performances for the tonal and three for the atonal piece) were transcribed from the voice recorder to manuscript paper and then marked according to pitch accuracy, rhythm accuracy, structural accuracy and fluency.[1] The pitch and rhythm scores were added up and created an overall 'technical accuracy score' for the piece; subsequently, all three scores, technical, structure and fluency, were converted in percentage scores and the resulting scores were added up using two different algorithms. For a detailed description of the scoring system see Chapter 2.3.3 and Appendix 4.

The 10-minute private practice sessions, two for each participant, were transcribed into descriptive text: the portions of music practised were listed, reporting repeats and any

---

[1] An explanation about the decision to produce hand-written transcriptions is given in section 2.2.1.5.

verbal comments; a sample transcription of a practice session of a non-solfège pianist is quoted below:

> Prima vista (30''). b.1 right hand 3t, both hands 4t, b.2 both hands 3t, b.1–2 3t, b. 4 both hands several t, b.1–4, b.4 several t, b.1–5 both hands, b.5 several t. (5') Continues using the same system to the end. Then plays b.1–8 several t. (attempts from memory?) – whenever a mistake occurs plays right hand separately for that bar and then joins both hands. Has played it over 3 t. start–finish at 10'.[2]

The same method of transcription was used for the two 2-minute recall sessions for each piece. The lack of visual element renders the transcriptions of these practice sessions somewhat incomplete in cases where the subject engaged in mental practice, but these were extremely few (see Chapter 3.3). For this study, the only finding used from the transcriptions of the practice sessions was the comparison between the strategies the subjects had actually used to memorise the music during the 10-minute practice session and the strategy they had stated they used in the questionnaire.

3.2. Preliminary analysis

3.2.1. Comparability of the pieces

The first fact that needed to be established in order to proceed with data analysis was that the memorised performances of the three tonal stimuli – T1, T2 and T3 – could be analysed as one single group; the same should be true for the three atonal stimuli A1, A2 and A3. The three pieces in each group had the same technical characteristics and had been piloted extensively, both by musicians serving as independent judges and by participants who actually practised and performed the pieces in an experimental setting in the pilot study (Chapter 2.2). In order to accommodate differences in instrumental and vocal range, the key of each piece varied across instruments; every instrumental and vocal version was screened by an independent professional musician and an

---

[2] A key to the transcription as well as some additional sample transcriptions are provided in Appendix 4.

independent student of the respective instrument or voice before it was given to participants in the experiment setting; alterations were made wherever necessary. Thus, the pieces that were finally used in the experiment were considered to be equivalent in terms of both technical and memorisation difficulty.

After the completion of the experiment and the marking of performances, the pieces were compared again in terms of the performance results they had yielded; in order to determine whether it would be plausible to pool the three tonal and the three atonal pieces in two unified groups. In order to establish this, the following null hypotheses were tested for both the tonal and the atonal group:

- Distribution and variance are equal across performance scores for T1, T2 and T3.
- Distribution and variance are equal across performance scores for A1, A2 and A3.

The above hypotheses were tested separately for the 1st, 2nd and 3rd memorised performances. Furthermore, data from the questionnaires regarding the pieces' level of difficulty were examined in order to estimate the comparability of the pieces according to the performers' own musical intuition.

3.2.1.1. Tonal stimuli

Normality tests for T1, T2 and T3 (Table 3.1) showed that for the 1st and 2nd performance, T2 and T3 scores had significantly non-normal distributions, whereas T1 appeared normal at $p > .05$. Data from the 3rd performance showed significant normality for both T1 and T3; it should be noted, however, that these data were considerably fewer due to many participants failing to attend their next-day session. Although these results suggested that T1 produced a different pattern of results, the common features of the distributions, such as prominent bimodality and negative skewness for all three pieces across performances, justified further investigation of the pieces' comparability. The difference, however, between T1 and the other two tonal pieces was noted and taken into account throughout the rest of the analyses; in most cases, two results will be

reported, the first comparing scores across the three pieces and the second comparing T1 scores with T2 and T3 scores grouped.

Tests of Normality

	Tonal piece tested	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
1st performance: Total score	T1	.121	30	.200[*]	.945	30	.123
	T2	.168	30	.030	.846	30	.001
	T3	.177	30	.017	.908	30	.013
2nd performance: Total score	T1	.122	30	.200[*]	.937	30	.077
	T2	.182	30	.013	.868	30	.002
	T3	.204	30	.003	.866	30	.001
3rd performance: Total score	T1	.154	21	.200[*]	.925	21	.109
	T2	.174	20	.114	.879	20	.017
	T3	.152	20	.200[*]	.920	20	.098

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Table 3.1: Normality tests for T1, T2 and T3 in performances 1, 2 and 3.

1st performance: The Kruskal-Wallis test for the distribution of scores of the 1st performance across the three pieces was found non-significant, $p > .05$, suggesting that the null hypothesis was valid and the piece performed did not have a significant effect on the performance score. Further grouping of the scores in (T1) and (T2, T3) did not change the results, as Mann-Whitney, Kolmogorov-Smirnov Z, Wald-Wolfowitz runs as well as the Kruskal-Wallis tests were all non-significant at $p > .05$, suggesting that the group in which the piece belonged did not have a significant effect on the distribution of the performance scores, either. Having divided the scores in two groups also made testing the variability of scores in each group possible: the Moses test of Extreme Reactions was applied and was found non-significant at $p > .05$, providing further support for the null hypothesis.

2nd performance: The Kruskal-Wallis test for the distribution of scores of the 2nd performance across the three pieces was found non-significant, $p > .05$, suggesting that the null hypothesis was valid. In testing the pieces in groups (T1 against T2,3), the

Mann-Whitney, Kolmogorov-Smirnov Z and the Kruskal-Wallis tests were non-significant at $p > .05$; only the Wald-Wolfowitz runs was found significant at $p < 0.05$, although the value was very close to the threshold of no significance ($p = 0.047$). The Moses Extreme Reactions test was also non-significant at $p > 0.5$, suggesting that the variance of scores was equal between the two groups.

3rd performance: The Kruskal-Wallis test for the distribution of scores of the 3rd performance across the three pieces was found non-significant, $p > .05$, suggesting that in the 3rd performance the piece performed did not have a significant effect on the performance score, either. Mann-Whitney, Kolmogorov-Smirnov Z, Wald-Wolfowitz runs as well as the Kruskal-Wallis tests all found that the distribution of performance scores was the same between T1 and T2,3 ($p > 0.5$); the Moses Extreme Reactions test was also non-significant ($p > 0.5$), supporting the null hypothesis that the piece group did not have a significant effect on the range of performance scores.

In order to seek further validation, MANOVA was also performed on the data for T1, T2 and T3, for all three performances, under the premise that the Pillai-Bartlett trace and, to a lesser extent, Wilks' lambda and Hotelling's trace are relatively robust to violations of multivariate normality and homogeneity of covariance matrices, when sample sizes are equal (Field, 2009). For our data, it has been established that the univariate normality condition has been violated, therefore multivariate normality is also violated; Box's test is significant ($p < .05$) showing that the homogeneity assumption has been violated as well, although variance according to piece performed for each performance separately is equal, $F(2, 58) = 0.04$, *ns*, for the 1st performance, $F(2, 58) = 0.02$, *ns*, for the 2nd performance and $F(2, 58) = 0.15$, *ns*, for the 3rd performance. Since, however, sample sizes are equal for all three pieces (N=30), the Pillai-Bartlett trace was used as tentative support for the comparability of the three pieces: using Pillai's trace, there was no significant effect of the piece performed on the scores of the 1st, 2nd and 3rd performance, $V = 0.18$, $F(6, 114) = 1.91$, $p > .05$.

Questionnaire data

There were two questions in the questionnaire, given to the participants during the 15-minute distraction interval, regarding the difficulty of the performed piece: the first asked them to rate the piece in terms of technical difficulty and the second in terms of memorisation difficulty. In both questions the participant was asked to select one of the following answers: a)Very difficult, b)Difficult, c)Neither difficult nor easy, d)Easy, e)Very easy. Participant responses are summarized in Tables 3.2 and 3.3:

Participant responses regarding the technical difficulty of the piece

Response / Piece tested	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
T1	0	2	7	13	8
T2	0	1	8	12	9
T3	1	4	7	13	5

Table 3.2: Participant responses regarding the tonal pieces' technical difficulty. N=30 for each piece.

Participant responses regarding the memorisation difficulty of the piece

Response / Piece tested	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
T1	0	12	13	5	0
T2	1	7	13	7	1
T3	1	8	14	6	0

Table 3.3: Participant responses regarding the tonal pieces' memorisation difficulty. N=30 for T1 and N=29 for T2 and T3, due to missing answers.

The three pieces elicited almost identical responses regarding technical difficulty, with the vast majority of the participants stating that the pieces were between average (neither difficult nor easy) and very easy. Memorisation difficulty ratings seem slightly more consistent with trends in performance scores, in that T1 tested slightly differently than T2 and T3, receiving the most responses for 'difficult' and the least responses for

'easy' or 'very easy'; this is also consistent with the fact that T1 produced the lowest mean scores across all three performances. The association, however, between the piece tested and the participant's response was found to be non-significant, $L\chi^2$ (8) = 5.68, p > .05.

Tonal Stimuli – Summary

The combination of analysis of the questionnaire data and the musical intuition of several musicians, who participated in the screening of the pieces, as well as the reports of the two independent judges, suggest that the three tonal pieces used in the experiment could be viewed as equivalent and analysed further as such. Statistical analysis of the performance scores generally supports this course of action; normality tests, however, give some cause for concern that that there could be a difference between performance results produced by T1 and the other two pieces, especially in the 1st and 2nd performance. For this reason, the analyses yielding statistically significant results using T1, T2 and T3 results grouped together will be double-checked with results from analyses using T1 results separately. The latter method has the obvious disadvantage of comparing very different sample sizes, since T1 results are exactly half of T2 and T3 results pooled together; under the circumstances, however, this comparison was considered preferable to disregarding the possible anomaly altogether and, in any case, results will be always looked at in context and taking into account each method's shortcomings.

3.2.1.2. Atonal stimuli

Normality tests for A1, A2 and A3 (Table 3.4) showed that for the 1st and 2nd performances all three pieces were significantly non-normal, p < 0.05, while for the 3rd performance A1 was found normal at p > 0.05, whereas A2 and A3 were found non-normal; again, it should be noted that the 3rd performance results were considerably fewer due to participants failing to attend the last session (next day) of the experiment. Normality was checked using the Shapiro-Wilk (S-W) test statistic rather than the Kolmogorov-Smirnov (K-S) in this case, where results from the two tests were in

conflict, for two reasons: the first was that the S-W is considered more powerful to deviations from normality (Field, 2009); the second was that visual representation of the data distribution supported the notion of non-normally distributed data: performance scores' distributions from all three pieces, across performances, had similar features, being notably platykurtic, negatively skewed and, in most cases, heavy-tailed.

Tests of Normality

	Atonal piece tested	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
1st performance:	A1	.120	34	.200[*]	.906	34	.007
Total score	A2	.190	29	.009	.871	29	.002
	A3	.161	29	.053	.897	29	.008
2nd performance:	A1	.142	34	.082	.922	34	.019
Total score	A2	.163	29	.048	.917	29	.026
	A3	.158	29	.063	.909	29	.016
3rd performance:	A1	.158	18	.200[*]	.898	18	.053
Total score	A2	.189	16	.129	.876	16	.033
	A3	.130	19	.200[*]	.899	19	.046

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Table 3.4: Normality tests for A1, A2 and A3 in performances 1, 2 and 3.

The Kruskal-Wallis test for the distribution of scores of the 1st performance of A1, A2 and A3 was found non-significant, $p > .05$, supporting the null hypothesis that the piece performed did not have a significant effect on the performance score. The null hypothesis was also supported for the 2nd and the 3rd performance of the three atonal pieces, suggesting that the three pieces yielded similar performance scores in all occasions.

Using Pillai's trace, there was no significant effect of piece tested on the 1st, 2nd and 3rd performance scores, $V = 0.08$, $F (6, 98) = 0.63$, $p > .05$; it should be noted that MANOVA tests were used only as tentative support, as the multivariate normality as well as the homogeneity assumptions were violated.

Questionnaire data

Participants received the same two questions for the atonal piece as for the tonal piece, regarding the technical and the memorisation difficulty of the performed piece, with the same selection of possible answers. Participant responses for the atonal pieces are summarized in Tables 3.5 and 3.6:

Participant responses regarding the technical difficulty of the piece

Response \\ Piece tested	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
A1	1	4	12	14	3
A2	2	5	6	11	5
A3	0	5	8	7	9

Table 3.5: Participant responses regarding the atonal pieces' technical difficulty. N=34 for A1 and N=29 for A2 and A3.

Participant responses regarding the memorisation difficulty of the piece

Response \\ Piece tested	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
A1	6	12	13	2	1
A2	5	15	7	1	1
A3	3	8	13	4	1

Table 36: Participant responses regarding the atonal pieces' memorisation difficulty. N=34 for A1 and N=29 for A2 and A3.

As in the tonal pieces, the three atonal pieces elicited almost identical responses regarding technical difficulty, with the vast majority of the participants stating that the pieces were between average (neither difficult nor easy) and very easy. Memorisation difficulty ratings were slightly more varied, with A2 featuring the least responses

between 'neither difficult nor easy' and 'very easy' and A3 the least responses between 'very difficult' and 'difficult', suggesting that participants may have found A2 the most challenging piece to memorise and A3 the easiest; this is also consistent with the fact that A2 yielded the lowest and A3 the highest mean scores in all three performances. The association, however, between the piece tested and the participant's response was found to be non significant, $L\chi^2$ (8) = 7.06, p > .05.

Atonal stimuli – Summary

The combination of statistical analysis of the performance scores and of the questionnaire data, as well as the musical intuition of the musicians who participated in the screening process and of the two independent judges, suggest that the three atonal pieces used in the experiment could be viewed as equivalent and analysed as such.

3.2.1.3. Comparability of the pieces: Summary

Analysis showed that T1, T2, T3 and A1, A2, A3 can be pooled together in a single Tonal and a single Atonal group respectively, so that their corresponding performance results are studied and analysed further in these groups; from now on every reference to the 'Tonal' or the 'Atonal' piece will imply any one of the three pieces included in the group, unless a specification is made. A possible reservation regarding T1 results was noted and will be taken in consideration in further analysis.

3.2.2. Comparability of different student groups

After establishing that performance scores from the three tonal and the three atonal pieces could be analysed in two unified groups, the comparability of student groups coming from different institutions needed to be checked. As described in section 3.1.2, a total of 93 participants were tested for the experiment: of those, 48 students came from countries in which solfège is taught and 45 students came from countries where solfège is not normally used.

The 48 solfège participants were further divided in three groups:
   I. 34 participants were tested in University 1a, country 1a, which was also their country of origin.
  II. 7 participants were tested in University 1b, country 1b, which was also their country of origin.
 III. 7 participants were tested in University 1c, country 1c; 5 of them came from 1c, and the other 2 came from different countries each.

The 45 non-solfège participants were further divided in three groups as well:
   I. 30 participants were tested in University 2a, country 2a: 28 of them came from 2a and the other 2 participants were from a different country.
  II. 11 participants were tested in University 2b, country 2b, which was also their country of origin.
 III. 4 participants were tested in University 1c, country 1c; 3 of them came from country 2c and 1 participant was from another country.

In summary, experiments were run in 5 different institutions: 3 were in solfège countries and 2 in non-solfège countries. Institution 1c was in a solfège country, but was used to test both solfège and non-solfège participants; the non-solfège participants tested in 1c were attending the institution on an exchange programme and had not attended any solfège classes. Other solfège participants tested in country 1c but not from 1c were regular students in the institution; the same was true for the two non-solfège participants in country 2a who did not come from 2a.

3.2.2.1. Comparison Groups

The comparison of participant groups was made according to the institution in which they were tested, for two main reasons:

    a. Numbers in different ethnic groups were too small to justify meaningful comparisons.

    b. Since all students tested in a single institution had been accepted and were attending that institution as students, it was considered plausible to accept that there were common features in those subjects' musical skills, both academic and practical.

Subsequently, the participants in the 3rd non-solfège group, who were tested in a solfège country (1c) but were exchange students from other countries, were put together in one group: 2c. Thus, the results of 6 different country groups were compared: 1a, 1b, 1c (solfège participants) and 2a, 2b and 2c (non-solfège participants). It is evident that certain groups in the comparison, especially 2c, are likely to be problematic in terms of homogeneity, as they include participants coming from different backgrounds; for this reason, the following analysis was necessary in order to determine whether or not the country group in itself affected performance scores.

3.2.2.2. Problems in the comparison

The most obvious problem concerning the comparison of the six different country groups was the large variation of sample sizes, as both the solfège and the non-solfège groups consist of three subgroups with a different number of participants: 34, 7 and 7 subjects for the solfège and 30, 11 and 4 subjects for the non-solfège group. This problem was a result of practical limitations of the study: students in different universities responded in very different and unpredictable ways to the call for participants and thus the recruitment of equal numbers of participants across universities was deemed impossible, mainly due to time restrictions.

The problem of disproportionate sample sizes was further exacerbated by a methodological flaw that arose during testing, regarding the proportion of AP possession within different country groups. AP possession was monitored by a relevant question in the questionnaire (see Appendix 3), to which there were three possible answers: 'Yes', 'No' and 'I don't know'. Out of the total 93 participants, 10 selected the third option; although answers in the questionnaires were not checked on the spot, 5 of the 10 participants who selected the 'I don't know' answer enquired about further clarifications regarding AP possession. As this happened during the 15-minute distraction task, in all 5 cases the researcher was prompted to explain the concept of AP and the conversation that ensued invariably led to certain conclusions regarding the particular participants' AP skills. These conclusions are, of course, only tentative and cannot compare to results from thorough testing for AP; on the other hand, the same is probably true for the self-assessment of AP skills the participants provided by answering the relevant question in the questionnaire. As such, this brief informal investigation led to 4 of the 5 participants being reported as 'unlikely to possess AP': these participants were unsure of the distinction between AP and good relative pitch and failed to identify with precision 5 random tones played for them on the piano. One of the participants was reported as 'likely to possess AP', as she identified with 100% precision 5 tones played for her on the piano and reported she could identify pitches produced from other instruments 'most of the times, but not always'. Absolute frequencies regarding AP possession for all 93 participants are provided in Table 3.7:

AP possession	TOTAL	NON-SOLFÈGE	SOLFÈGE
Non-AP	75	37	38
AP	9	3	6
Unsure	5	2	3
Likely AP	1	0	1
Unlikely AP	4	3	1

Table 3.7: Participant data regarding AP possession, derived from questionnaires and informal testing. Participants who answered 'No' to the question 'Do you have absolute pitch?' are listed as non-AP; participants who answered 'Yes' are listed as AP and those who answered 'I don't know' are listed as Unsure, Likely AP or Unlikely AP, depending on whether or not they sought further clarifications from the researcher regarding AP.

If the groups of AP and Likely AP participants are pooled together and the same is done for the non-AP and Unlikely AP groups, an imbalance in AP possessors is observed between the solfège and the non-solfège groups: in the solfège group there is a total of 7 AP participants (14.58%), against only 3 AP possessors (6.38%) in the non-solfège group. The occurrence of AP in the general population is normally around 1 in 10,000 (Ward, 1999), although this number increases to 1 in 1,500 amongst amateur student musicians (Profita and Bidder, 1998) and is highest amongst professional musicians, reaching up to 15% (Baharloo et al, 1998; Gregersen et al, 1999; Gregersen et al, 2000). According to these facts, the occurrence of 14.58% in the solfège group, although still high, especially in relation to the non-solfège group, is within previously reported levels. The effect of AP possession on the quality of the memorised performance is going to be investigated separately in the following sections; what needs to be examined further at this point is the distribution of AP participants in solfège countries. Country 1a included 3 AP participants, accounting for 8.82% of the total number of participants from 1a; country 1c included 1 AP participant, accounting for 14.29% of the total amount of 1c participants and country 1b included 3 AP participants, accounting for 42.86% of the total amount of 1c participants. Even when taking the small total number of subjects from 1c into consideration, the percentage is extremely high; the fact that 3 out of 7 participants from the same country happened to be AP possessors was considered extremely unlikely to be an outcome of random sampling procedures.

Once again, this was believed to be a result of a practical issue, which could not have been foreseen or avoided: all collaborating members of staff, in all 6 countries, had been specifically asked to approach the students who thought would be most willing to participate in the research, regardless of aural or performance skills level; the researcher was entirely responsible for the final screening of participants (for the selection criteria for participants see Chapter 3.1.2). In all cases but the one of 1b, the teachers involved in participant recruitment acted as instructed and selected subjects based on their interest in participating, reliability and punctuality; the lecturer in 1b, however, who was also responsible for teaching aural skills in that University, apparently selected students based on their excellence in aural and performance skills, choosing the best students. This was supported both from the high occurrence of AP in the group (3 out of 7 students) and from the fact that a large percentage of the students approached (5 out

of 11 students) failed either to show up or even set an appointment for the experiment; this suggested that reliability and desire to participate in the research was unlikely to have been a crucial selection criterion.

Despite the fact that AP possession does not necessarily affect performance in music memorisation tasks, the inequality of AP subjects between the solfège and the non-solfège groups was taken into consideration in the following analyses. The easiest solution in order to prevent a potential bias in the results would be to analyse separately the data from the 10 AP participants from both groups, throughout the study; the elimination of AP subjects, however, would result in the further diminution of group 1b to 4 participants, which would augment the already prominent problem of unequal sample sizes. For this reason, AP data were analysed both separately and in conjunction with non-AP data for the countries' comparison and the results were compared in order to determine the overall significance of AP possession in memorised performance for subsequent analyses.

3.2.2.3. Comparison hypotheses

In order to establish the comparability of the country groups, the following null hypotheses needed to be tested:

Distribution and variance are equal across performance scores for the following country groups, performances, and pieces:

- 1a, 1b, 1c – Tonal Piece, 1st performance
- 1a, 1b, 1c – Tonal Piece, 2nd performance
- 1a, 1b, 1c – Atonal Piece, 1st performance
- 1a, 1b, 1c – Atonal Piece, 2nd performance
- 2a, 2b, 2c – Tonal Piece, 1st performance
- 2a, 2b, 2c – Tonal Piece, 2nd performance
- 2a, 2b, 2c – Atonal Piece, 1st performance
- 2a, 2b, 2c – Atonal Piece, 2nd performance

Performance results from the 3rd performance were not used in this analysis due to the large number of missing cases; if the null hypotheses were supported for the first two performances, it was considered it would hold enough proof that the 6 country groups were similar.

3.2.2.4. Comparison results: Tonal pieces

Shapiro-Wilk test statistics showed that the assumption of normality within country and AP possession groups had been violated for most subgroups in the 1st and 2nd performance of the Tonal Piece, so non-parametric tests were selected to compare the distributions and variances across these groups. The Kruskal-Wallis test statistic was found non-significant at $p > .05$ for all the solfège and the non-solfège country groups in the 1st and 2nd performances of the Tonal Piece, supporting the null hypothesis that the distribution of memorised performance scores was the same across groups. Multiple regression analysis was also carried out in both the solfège and the non-solfège groups in order to seek further support for the null hypothesis.

Solfège countries: 1a, 1b, 1c

Multiple regression analysis was carried for the 1st and 2nd performance scores, using the change of country and AP possession as predictors in the same block. As the residuals plots showed cause for concern regarding heteroscedasticity, a robust version of multiple regression was used, applying the bootstrap method. Results are shown in Tables 3.8 and 3.9:

Tonal piece, 1ˢᵗ performance: Solfège group

	B	SE B[a]	β
Constant	604.58	49.078	
1a vs. 1b	- 120.51	104.877	- .15*
1a vs. 1c	178.23	106.007	.24*
Non-AP vs. AP	58.91	148.073	.068*
Non-AP vs. unsure	216.96	109.210	.25**

Note: $R^2 = .14$, $p > .05$
* $p > .05$ (2-tailed), *ns*
** $p < .05$ (2-tailed), *s*

a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.8: Multiple regression statistics for the 1ˢᵗ memorised performance of the Tonal Piece by the solfège groups. For country groups, country 1a was chosen as the base predictor to which the other 2 were compared, as it was the country with the largest number of participants. For AP possession groups, the non-AP group was compared to AP possessors and participants who had selected the 'I don't know' answer.

Tonal piece, 2nd performance: Solfège group

	B	SE B[a]	β
Constant	698.33	45.57	
1a vs. 1b	- 317.85	115.10	- .38**
1a vs. 1c	14.95	182.38	.02*
Non-AP vs. AP	- 50.81	202.94	- .06*
Non-AP vs. unsure	219.35	119.43	.24**

Note: $R^2 = .16$, $p > .05$
* $p > .05$ (2-tailed), *ns*

** $p < .05$ (2-tailed), *s*

a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.9: Multiple regression statistics for the 2ⁿᵈ memorised performance of the Tonal Piece by the solfège groups. The same baseline predictors were chosen for comparisons between country groups and AP possessor groups.

The overall effect of the predictors on performance scores appeared to be non-significant for both the 1st and the 2nd performance of the Tonal Piece ($R^2 = .14$ for the 1st and $R^2 = .16$ for the 2nd performance, both at $p > .05$, *ns*). Within this non-significant result, however, the standardized β values of the AP groups seem to be significant: belonging in the 'Uncertain' group (participants who answered they did not know whether they had AP or not) seemed to affect performance results in both the 1st and the 2nd memorised performance. Considering that the absolute number of participants belonging in the 'Uncertain' group was just 3 for the solfège group (see Table 3.7) and that, for the 1st performance, the bootstrapped confidence interval boundaries for the Uncertain group's β crossed zero, this finding was attributed to a Type I error and no action was taken to rectify the situation. A significant β value was also found for the difference between countries 1a and 1b, but only for the 2nd performance: this, too, was rejected as a Type I error, given the non-significant value of R^2 and the fact that this effect did not appear in both performances. If the significant value of β between 1a and 1b had been accompanied by a significant value of β for the difference between non-AP and AP groups, it would have raised some concern, in relation to the aforementioned disproportionate AP occurrence in 1b; this, however, was not the case so 1b subjects' performance did not appear affected by AP possession, compared to 1a.

Non-solfège countries: 2a, 2b, 2c

Multiple regression analysis was carried for the 1st and 2nd performance scores of the non-solfège group, again using the change of country and AP possession as predictors in the same block. The residuals plots showed to violate the assumption of homoscedasticity in this case too, so the bootstrap method was used again. Results are shown in Tables 3.10 and 3.11:

Tonal piece, 1st performance: Non-solfège group

	B	SE B[a]	β
Constant	629.39	61.69	
2a vs. 2b	104.99	90.81	.16*
2a vs. 2c	123.24	142.74	.13*
Non-AP vs. AP	241.30	121.94	.22**
Non-AP vs. unsure	- 78.78	116.24	- .09*

Note: $R^2 = .09$, $p > .05$
* $p > .05$ (2-tailed), *ns*
** $p < .05$ (2-tailed), *s*

a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.10: Multiple regression statistics for the 1st memorised performance of the Tonal Piece by the non-solfège groups. For country groups, country 2a was chosen as the base predictor to which the other 2 were compared, as it was the country with the largest number of participants. For AP possession groups, the non-AP group was compared to AP possessors and participants who had selected the 'I don't know' answer.

Tonal piece, 2nd performance: Non-solfège group

	B	SE B[a]	β
Constant	676.44	55.25	
2a vs. 2b	7.71	90.31	.01*
2a vs. 2c	76.92	129.14	.09*
Non-AP vs. AP	217.99	125.93	.21**
Non-AP vs. unsure	- 54.14	103.72	- .07**

Note: $R^2 = .07$, $p > .05$
* $p > .05$ (2-tailed), *ns*
** $p < .05$ (2-tailed), *s*

a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.11: Multiple regression statistics for the 2nd memorised performance of the Tonal Piece by the non-solfège groups. The same baseline predictors were chosen for comparisons between country groups and AP possessor groups.

As demonstrated in the non-significant values of R^2 in Tables 3.10 and 3.11 ($R^2 = .09$ for the 1st and $R^2 = .07$ for the 2nd performance, both at $p > .05$, *ns*) country and AP group are not reliable predictors for memorised performance scores. Within these non-significant results, AP possession produced significant β-values for both performances: when participants belong to the AP group, their performance result was more likely to be different than when they belong to the non-AP or than when they did not know whether or not they possess AP. For both performances, however, the confidence intervals for the difference between the AP and the non-AP group crossed zero; confidence intervals along with the discordance with R^2 values led to the conclusion that, as in the case of the solfège group, this finding could also be attributed to a Type I error.

General summary

The changes in R^2 values for both performances were non-significant in both the solfège and the non-solfège groups, indicating that the overall effect of the predictors of country and AP on the score was extremely small. Findings suggesting that AP possessors performed differently than non-AP possessors in the non-solfège group were attributed to a Type I error.

3.2.2.5. Comparison results: Atonal pieces

Shapiro-Wilk test statistics showed that the assumption of normality within country and AP possession groups had been violated for most subgroups in the 1st and 2nd performance of the Atonal Piece, so non-parametric tests were selected to compare the distributions and variances across these groups. The independent samples Kruskal-Wallis test statistics were found non significant at $p > .05$ for both the solfège and the non-solfège country groups, supporting the null hypothesis that the distribution of performance scores was the same across groups. Multiple regression analysis was also carried out separately for the solfège and the non-solfège group, in order to seek further validation for the null hypothesis.

Solfège group: 1a, 1b, 1c

The plot of standardized residuals against standardized predicted values suggested that the homoscedasticity assumption had been violated, therefore multiple regression was carried out on bootstrap samples. Results are shown in Tables 3.12 and 3.13:

Atonal piece, 1st performance: Solfège group

	B	SE B^a	β
Constant	668.40	42.34	
1a vs. 1b	- 28.78	110.53	- .04*
1a vs. 1c	111.35	104.18	.17*
Non-AP vs. AP	127.92	131.84	.15*
Non-AP vs. unsure	130.95	118.87	.17*

Note: $R^2 = .09$, $p > .05$
* $p > .05$ (2-tailed), *ns*
a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.12: Multiple regression statistics for the 1st memorised performance of the Atonal Piece by the solfège groups. Comparison groups are the same as in the regression model for the performance results of the Tonal Piece.

Atonal piece, 2nd performance: Solfège group

	B	SE B^a	β
Constant	686.31	43.82	
1a vs. 1b	- 164.33	130.88	- .23*
1a vs. 1c	24.33	99.85	.04*
Non-AP vs. AP	- 88.86	150.08	- .10*
Non-AP vs. unsure	208.38	95.05	.27**

Note: $R^2 = .10$, $p > .05$
* $p > .05$ (2-tailed), *ns*
** $p < .05$ (2-tailed), *s*
a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.13: Multiple regression statistics for the 2nd memorised performance of the Atonal Piece by the solfège groups.

The values of R^2 for both performances are non significant, which supported the null hypothesis that neither country nor AP possession group had a significant effect on memorised performance scores. Within this non-significant result, participants who had selected the 'I don't know' response in the relevant AP question, seemed to produce significantly different results than their non-AP and their AP counterparts; considering the non-significant R^2 value, this finding was attributed to a Type I error.

Non-solfège countries: 2a, 2b, 2c

Multiple regression analysis was carried for the 1st and 2nd performance scores of the non-solfège group, again using the change of country and AP possession as predictors in the same block. The residuals plots showed to violate the assumption of homoscedasticity in this case too, so the bootstrap method was used again. Results are shown in Tables 3.14 and 3.15:

Atonal piece, 1st performance: Non-solfège group

	B	SE B[a]	β
Constant	754.94	50.16	
2a vs. 2b	- 34.64	91.23	- .062*
2a vs. 2c	23.43	111.18	.028*
Non-AP vs. AP	212.91	70.72	.22**
Non-AP vs. unsure	28.71	95.66	.04*

Note: $R^2 = .06$, $p > .05$
* $p > .05$ (2-tailed), *ns*
** $p < .05$ (2-tailed), *s*
a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.14: Multiple regression statistics for the 1st memorised performance of the Atonal Piece by the non-solfège groups.

Atonal piece, 2nd performance: Non-solfège group

	B	**SE B**[a]	**β**
Constant	716.59	47.14	
2a vs. 2b	-61.12	78.00	-.11*
2a vs. 2c	51.12	98.16	.062*
Non-AP vs. AP	222.93	81.81	.24**
Non-AP vs. unsure	53.76	107.10	.07*

Note: $R^2 = .09$, p > .05
* p > .05 (2-tailed), *ns*
a. Bootstrapped values. Two-tailed significance for β is also based on bootstrap results.

Table 3.15: Multiple regression statistics for the 2nd memorised performance of the Atonal Piece by the non-solfège groups.

R^2 values for both performances of the Atonal Piece by the non-solfège groups are extremely low and non-significant, providing further support for the null hypothesis that different country groups and AP possession did not have an effect on memorisation performance scores. Within this non-significant result, AP possession seems to produce a significant effect on the scores of both the 1st and the 2nd performance of the Atonal Piece; as in the 2nd performance confidence interval boundaries crossed zero and the significant findings were not consistent with R^2 values, they were attributed to a Type I error.

3.2.2.6. Comparability of student groups: Summary

Non-parametric tests as well as multiple regression models tested for the 1st and the 2nd memorised performances of both the Tonal and the Atonal Pieces suggested that changes in country and AP possession did not have a significant effect on performance. Despite the fact, however, that test statistics did not suggest that AP possession had a significant effect on performance, the exclusion of AP participants from the dataset was decided for the rest of the analysis. This decision was taken in order to eliminate potential interactions between AP possession and other factors, which could alter the results in unpredictable ways; the imbalance between AP subject numbers in the solfège

and the non-solfège group also functioned as an argument in favour of the exclusion of AP participants.

The uniqueness of AP possession would justify more extensive, separate analysis of the results of the AP group in the study: further examination should aim to determine the details of AP memorisation performance and the ties between AP possession and other predictors, which altogether presumably constitute a unique profile of music memorisers. The limited size of 10 AP subjects in the current sample, however, did not allow this type of analysis to be carried out within the framework of the present study; a larger sample and more target-specific testing are required for an authoritative, meaningful study of AP music memorisation. On the other hand, the 'Unsure' group did not produce a significantly different pattern of results from the non-AP group, apart from of the 2nd performances of both the Tonal and the Atonal Pieces by the solfège group, which was attributed to a Type I error; as the total number of participants who selected the 'I don't know' answer in the AP question was very small and these participants were distributed evenly between the solfège and the non-solfège group, data from these participants were included in further analyses.

3.3. Main analysis

The major objective of data analysis was to reveal the factors that affected memorised performance results the most. Having eliminated the type of piece (T1, T2, T3 or A1, A2, A3 respectively) and the country of the subjects as possible factors that crucially affect performance, the following three parameters were examined:

1. **Instrument type.** The participants were divided in four different categories: *i) vocalists*, including male and female first-study singers, *ii) string players*, including violin, viola, violoncello and double bass players, *iii) wind players*, including both woodwind and brass instrument players and *iv) pianists*.

2. **Solfège use.** This was examined in two levels: in the first level, subjects were divided into solfège and non-solfège groups according to whether they came from a solfège country and therefore had been asked to use solfège during the

experiment. In the second level, participants in the solfège group were filtered further according to the data from the 10-minute practice transcriptions: participants who had *actually* used solfège were separated from those who had not used solfège in an audible way during practice. This categorisation was useful in distinguishing between the lower, or first, level of acquaintance with solfège note names, which does not necessarily imply use of these elements in musical reality, and the higher, or second, level of extended and cultivated familiarity with solfège labels, which is presupposed for use of solfège in music practice. The distinction between the two levels may also be perceived as a potential distinction between the ***implicit*** and the ***explicit*** effects of solfège knowledge. For this reason, the labels *Solfège use (I)* and *Solfège use (E)* are going to be used hereon to denote whether the first-level (implicit) or the second-level (explicit) solfège knowledge is being compared to the non-solfège level.

3. **Memorisation frequency.** This parameter was based on responses to the question 'How often do you memorise music for performance?' in the participants' questionnaire (see Appendix 3). Possible answers were: a) Always/Almost always, b) Often, c) Occasionally and d) Never/Almost never.

Initially, a 4x4x2 MANOVA was run, using the results of the 1st performance for the Tonal and the Atonal Piece as the two dependent variables. Instrument type, Solfège use (Implicit vs. non-solfège) and Memorisation frequency were the independent variables. The assumption of multivariate normality had been violated and the same was true for the homogeneity of covariance matrices; as the latter was measured using Box's test statistic, however, the significant result may have reflected the test's sensitivity to non-normality. Nevertheless, the violation of assumptions was taken into consideration when interpreting the results.

Using Pillai-Bartlett's trace, there was a significant effect of Instrument type on 1st Performance scores for the Tonal and Atonal pieces, $V = 0.26$, $F(6, 106) = 2.66$, $p < .05$, while the effects of Memorisation frequency and Solfège Use (I) were found non-significant at $p > .05$. The Pillai-Bartlett trace has been proven to be relatively robust to violations of normality when sample sizes are equal; Instrument type subgroups, however, do not contain equal participant numbers: after the elimination of AP

participants, there were 30 singers, 25 pianists, 13 wind and 12 string players. Tabachnick and Fidell (2007, quoted in Field, 2009, p. 604) report that probability values are conservative, and can therefore be trusted, when the larger samples have produced larger variances. This is the case for variances between Instrument type subgroups, so the significance of the Pillai-Bartlett trace was accepted as reliable. On the other hand, the Pillai-Bartlett trace for Memorisation frequency and for Solfège use (I) on the 1st performance of the Tonal and the Atonal piece was found non-significant; as for both variables variance grew with subgroup size and therefore the test statistic should be considered conservative, the effect of Memorisation frequency and Solfège use on performance was examined further.

Multiple regression analysis was conducted on 1st performance results of the Tonal piece, using Instrument type as a predictor variable in the first block and Solfège use (E) with Memorisation Frequency in the second block; residuals plots showed signs of heteroscedasticity, so bootstrap samples were used. In addition to using explicit Solfège use in place of the implicit, which had been used in the previous MANOVA, Memorisation frequency subgroups were reduced to two: participants who had responded they memorise music for performance a) Always or b) Often were pooled together and the same was done for participants who had responded either c) Occasionally or d) Never, to the same question. The reason for these changes in the predictor variables was the desire to maximize any traceable effects; subjects who had audibly (explicitly) used solfège during the 10-minute practice were thought to be more likely to have displayed a solfège-related effect, whereas the creation of a single high-frequency and a single low-frequency memorisation group was aiming to stress the potential differences between groups, which would possibly have been less prominent in the 4-group situation.

Following the MANOVA test statistics, the hypotheses tested by the regression model were directional: subjects performing the task on a string or wind instrument were expected to score higher than singers and singers were expected to perform better than pianists. Directional hypotheses were also formulated for Memorisation frequency and Solfège use: participants in the high Memorisation frequency group were expected to perform better than participants in the low-frequency group, and participants who had

used solfège explicitly were expected to perform better than participants who had not used solfège. Multiple regression statistics are shown in Table 3.16:

Tonal piece, 1ˢᵗ performance

	B	SE Bᵃ	β
Step 1			
Constant	549.30	48.68	
Singers vs. Strings	191.75	100.42	.25**
Singers vs. Winds	319.61	71.48	.43***
Singers vs. Pianists	39.07	67.10	.07*
Step 2			
Constant	417.31	70.17	
Singers vs. Strings	253.15	93.54	.33**
Singers vs. Winds	372.53	70.32	.50***
Singers vs. Pianists	115.48	70.53	.10**
Solfège use (E)	28.66	59.37	.05*
Memorisation frequency (2 groups)	152.23	54.48	.28**

$R^2 = .18$, p = .001 for Step 1, $\Delta R^2 = .06$, p < .05 for Step 2.

* p > .05, *ns*
** p < .05, *s*
*** p ≤ .01, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.16: Multiple regression statistics for the 1ˢᵗ memorised performance of the Tonal Piece. Singers were chosen as the base group to which other instrument types were compared; the other two predictors contain only two categories: no solfège used and solfège used, or Low and High Memorisation frequency, respectively.

The overall effect of predictors on performance appears significant at p < .05; the values of R^2, however, although significant, are still relatively low: $R^2 = .18$ for Step 1 and $R^2 = .25$ for Step 2, which leaves a 75% of the change in performance scores unexplained by changes in Instrument type, Memorisation frequency, or Solfège use (E). Nevertheless,

significant beta values indicate that, for the 1st performance of the Tonal piece, the biggest change in scores appeared between singers and wind instrument players and singers and string players; the difference produced by the pianist group only appeared significant in Step 2 of the model. Since inclusion in the pianist group did not have a high correlation with another predictor variable, the fact that this group produced a significant change in the model only during Step 2 was interpreted as a result of the pianist group being borderline significant compared to the others; the value of p = .049, confirmed this suspicion. It is important to note that the hypothesis that pianists would perform worse than singers was not supported; inclusion in the pianist group appeared to have an important *positive* effect on performance scores (β = .10, p < .05). High Memorisation frequency also seems to play an important role in the shaping of memorised performance scores, although less important than belonging in either the string player or the wind player group. The effect of Instrument type was corroborated by the regression model for the 2nd performance of the Tonal piece; results are shown in Table 3.17:

Tonal piece, 2nd performance

	B	SE B[a]	β
Step 1			
Constant	578.63	50.81	
Singers vs. Strings	233.63	82.07	.32**
Singers vs. Winds	286.33	78.79	.40**
Singers vs. Pianists	65.82	66.52	.12*
Step 2			
Constant	496.39	71.26	
Singers vs. Strings	271.55	83.42	.37**
Singers vs. Winds	319.19	78.77	.45***
Singers vs. Pianists	113.57	69.40	.20*
Solfège use (E)	9.45	58.33	.02*
Memorisation frequency (2 groups)	100.70	55.79	.19**

Note: $R^2 = .18$, $p < .05$ for Step 1, $\Delta R^2 = .03$, $p > .05$ for Step 2.

* $p > .05$, *ns*
** $p < .05$, *s*
*** $p \leq .01$, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.17: Multiple regression statistics for the 2nd performance of the Tonal Piece. Predictor categories were kept the same as in the regression model for the 1st performance.

The difference between singers and the string and wind player groups is, once more, the most significant predictor of the model, as demonstrated both by significant beta values for these groups as well as the significance of $\Delta R^2 = .18$ in the first step of the model, where only Instrument type had been entered as predictor. The difference between pianists and singers was found non-significant for the 2nd performance. Solfège use (E) does not seem to affect performance, in line with findings for the 1st performance; Memorisation frequency, however, appears to have a significant effect, as it did for the 1st performance results.

Multiple regression models were also applied in the 1st and 2nd performances of the Atonal piece: for these analyses, too, hierarchy for predictors was kept the same as for the Tonal piece, in order to be able to compare results between the two pieces; bootstrap samples were also used in all occasions for the same reason, although the application of robust methods was also dictated by violations of the homoscedasticity assumption. Results for the Atonal piece performances are given in Tables 3.18 and 3.19:

Atonal piece, 1st performance

	B	SE B[a]	β
Step 1			
Constant	616.88	46.64	
Singers vs. Strings	253.548	62.63	.37***
Singers vs. Winds	312.978	53.67	.47***
Singers vs. Pianists	46.348	62.37	.09*
Step 2			
Constant	596.448	58.09	
Singers vs. Strings	261.448	62.81	.38***
Singers vs. Winds	320.578	54.72	.49***
Singers vs. Pianists	59.10	64.78	.11*
Solfège use (E)	- 34.12	53.71	- .07*
Memorisation frequency (2 groups)	50.57	41.66	.11*

Note: $R^2 = .26$, $p < .001$ for Step 1, $\Delta R^2 = .02$, $p > .05$ for Step 2.

* $p > .05$, *ns*
** $p < .05$, *s*
*** $p \leq .01$, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.18: Multiple regression statistics for the 1st memorised performance of the Atonal Piece. Note that the change in the model is only significant during Step 1, where Instrument type was the predictor variable.

Atonal piece, 2nd performance

	B	SE B[a]	β
Step 1			
Constant	598.20	44.10	
Singers vs. Strings	266.32	73.76	.39***
Singers vs. Winds	269.48	63.08	.41***
Singers vs. Pianists	49.16	58.05	.10*
Step 2			
Constant	598.57	60.19	
Singers vs. Strings	264.94	76.26	.39***
Singers vs. Winds	268.81	68.36	.41***
Singers vs. Pianists	49.28	66.87	.10*
Solfège use (E)	- 26.86	57.49	- .05*
Memorisation frequency (2 groups)	18.55	48.89	.04*

Note: $R^2 = .23$, $p < .001$ for Step 1, $\Delta R^2 = .004$, $p > .05$ for Step 2.

* $p > .05$, *ns*
** $p < .05$, *s*
*** $p \leq .01$, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.19: Multiple regression statistics for the 2nd performance of the Atonal piece. Note that, as in the 1st performance of the Atonal piece (Table 3.18), the change in the model is only significant in Step 1.

ΔR^2 changes in the models for the 1st and 2nd performance of the Atonal piece are only significant for Step 1, indicating that Instrument type was the only crucial predictor for performance, for the Atonal piece. This is confirmed by the non-significant beta values for changes in Memorisation frequency and Solfège use for both performances. The difference between the singer and the pianist groups, which appeared ambivalent in performances of the Tonal piece, appeared non-significant in both performances for the Atonal piece.

After examining Instrument type, Memorisation frequency and Solfège use effect on performance scores and finding that Instrument type and Memorisation frequency could account for up to 26% of change in the regression model, the most imperative research objective was to find out if there were other dependent variables which had been measured and had a significant effect on performance scores. After an inspection of all the variables measured in the questionnaire, multiple regression analysis was run on the 1st and 2nd performance of the Tonal piece, using the following predictors:

1. **Instrument type.** This variable had been found significant in the previous models; since string and wind players appeared to be the group that produced the greatest changes in scores compared to singers, strings and winds were placed in a single group, under the label 'monophonic instruments'. The existence of this group was supported by findings and was also part of the original experimental design: participants had been selected, amongst other screening factors, on the basis of whether they were vocalists, pianists, or monophonic instrument players.

2. **Memorisation frequency (2 groups).** This variable had also been found significant in the previous analyses; the 2-group division was kept instead of the original 4-level frequency division resulting directly from questionnaire responses.

3. **Use of a memorisation strategy (general).** Participants' responses in the question: 'Do you employ any particular strategies for memorising music?' were used to divide subjects in two groups: participants who had responded 'Yes' were placed in one group and participants who had responded 'No' or 'I don't know' were placed the second group.

4. **Use of a memorisation strategy (specific).** Participant responses in the question about particular strategies they use to memorise music, such as repeating, chunking, analysis and others (see Appendix 3) were cross-validated with transcriptions of the 10-minute practice sessions. In many occasions, transcription records revealed a completely different approach to the piece than the one described in the questionnaire response; reasons for this will be analysed in the Discussion section. For the regression analysis, participants were categorized according to whether they had *actually* used the techniques they had asserted they normally use. The inclusion of this variable should not be regarded, in conjunction with the previous predictor, as an attempt to double the effect of memorisation strategy: while the specific memorisation strategy actually employed by the

subject was a measure of what they *did* regarding the piece, the general strategy they had in mind reflected the more abstract measure of how they *thought about* the piece, in terms of their preconceptions about how they should form their approach to it. According to questionnaire responses, five different memorisation strategies were determined, with 'no strategy' being used as the baseline category to which the others were compared: i) Repeating the piece many times from the beginning to the end, ii) Dividing the piece in smaller chunks and learning them separately, iii) Determining structural elements of the piece, iv) Mental practice of the piece, without using the instrument, v) Other strategies.

Regression statistics for the 1st and 2nd performance of the Tonal piece, using Instrument type, Memorisation frequency, Memorisation strategy (general), and Memorisation strategy (specific)   as predictors, are given in Table 3.20:

Tonal piece	1st performance			2nd performance		
	B	SE B[a]	β	B	SE B[a]	β
Step 1						
Constant	741.38	57.75		795.57	51.11	
Monophonic vs. Singers	-307.36	61.24	-.54***	-293.78	66.79	-.54***
Monophonic vs. Pianists	-201.12	66.98	-.34**	-183.20	62.10	-.32**
Memorisation frequency (2 groups)	150.37	52.33	.27**	100.22	52.88	.19**
Step 2						
Constant	571.69	70.21		615.65	67.91	
Monophonic vs. Singers	- 257.08	68.50	- .45***	- 255.05	70.80	- .47***
Monophonic vs. Pianists	- 183.21	69.11	- .31**	- 170.56	62.99	- .30**
Memorisation frequency (2 groups)	108.34	52.16	.20**	69.34	53.31	.13*
Memorisation strategies (2 groups)	153.81	66.96	.19**	163.80	72.58	.21**
No strategy vs. repeat	6.06	74.39	.01*	- 1.13	68.85	- .00*
No strategy vs. chunking	10.89	70.98	.02*	38.81	65.75	.07*
No strategy vs. analysis	231.74	75.17	.27**	184.30	75.47	.22**
No strategy vs. mental practice	57.77	138.39	.05*	82.58	174.51	.08*
No strategy vs. other	3.44	216.02	.00*	- 68.35	235.68	- .06*

1st performance: $R^2 = .23$, p < .001 for Step 1, $\Delta R^2 = .13$, p < .05 for Step 2.
2nd performance: $R^2 = .21$, p = .001 for Step 1, $\Delta R^2 = .13$, p < .05 for Step 2.
* p > .05, *ns*
** p < .05, *s*
*** p ≤ .01, *s*
   a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.20: Multiple regression statistics for the 1st and 2nd performance of the Tonal piece.

As shown by ΔR^2 values, the inclusion of these predictors improved the model significantly. In both performances, monophonic instruments appeared to do significantly better than singers and pianists; as the previous analysis had shown that pianists performed better than singers, the conclusion was drawn that Instrument type was important in predicting performance in the Tonal piece, with monophonic instrument players scoring highest and singers scoring the lowest. Memorisation frequency appeared to have a significant effect on the results of the 1st performance, but this effect diminished in the 2nd performance; the difference in scores of the 1st and the 2nd performances will be examined in a subsequent section.

The use of a particular memorisation strategy, versus the absence of strategy, also appeared to have a significant effect on scores, although the only specific strategy which appeared to make a difference, amongst the ones actually used on task, was the one involving the structural analysis of the piece. The overall impact of the predictors appeared greater for the 1st than for the 2nd performance of the Tonal piece.

Apart from the predictors used in multiple regression for the Tonal piece – Instrument type, Memorisation frequency, Memorisation difficulty and Memorisation strategy – another two predictors were used for the Atonal piece:

1. **Frequency of atonal music practice.** This variable signified the frequency with which subjects used to practise or perform atonal music, as denoted by their responses in the relevant question of the questionnaire (see Appendix 3). As in the memorisation frequency variable, participants were divided in two frequency groups.

2. **Familiarity with atonal music.** In the relevant question from the questionnaire (Appendix 3), participants were able to choose between five possible responses: a) Extremely familiar, b) Familiar, c) Neither familiar nor unfamiliar, d) Unfamiliar, e) Extremely unfamiliar. Participants who had responded c) Neither familiar nor unfamiliar were chosen as the baseline group to which the other participants were compared: subjects who had answered a) or b) were placed in the 'familiar' group, whereas subjects who had responded d) or e) were placed in the 'unfamiliar' group.

All the above variables were entered in a multiple regression model for the 1st and 2nd performances of the Atonal piece, which yielded non-significant results. This led to a reduction of the predictors, leading to the results displayed in Table 3.21:

Atonal piece	1st performance			2nd performance		
	B	SE B[a]	β	B	SE B[a]	β
Constant	880.17	61.34		815.34	68.64	
Monophonic vs. Singers	- 295.51	60.03	- .59***	- 273.69	63.70	- .56***
Monophonic vs. Pianists	- 251.00	55.96	- .48***	- 237.79	55.44	- .46***
Frequency atonal (2 groups)	- 84.05	94.95	- .13*	- 70.36	94.34	- .11*
Average vs. unfamiliar	90.51	84.45	.16*	47.11	79.56	.08*
Average vs. familiar	- 19.56	59.88	- .04*	4.36	57.67	.01*
No strategy vs. repeat	55.91	66.56	.08*	12.13	77.76	.02*
No strategy vs. chunking	37.62	46.79	.08*	116.24	47.04	.24**
No strategy vs. analysis	27.67	102.16	.04*	46.10	94.13	.06*
No strategy vs. mental practice	- 34.56	98.37	- .04*	- 98.90	117.11	- .11*
No strategy vs. other	116.24	149.18	.09*	130.49	142.24	.10*

1st performance: $R^2 = .33$, p = .001.
2nd performance: $R^2 = .33$, p = .001.
p > .05, *ns*; ** p < .05, *s*; *** p < .01, *s*.
a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 3.21: Multiple regression statistics for the 1st and 2nd performance of the Atonal Piece.

R^2 values show that the predictors significantly affected the model for both performances. Regarding individual predictors, the most interesting result was the standardised beta value for the difference generated by the group of participants who had used chunking as their memorisation strategy: this was found significant, but only for the 2nd performance. As the beta values were not similar for the two performances, which could have justified a borderline significance for the 2nd performance, or a borderline non-significance for the 1st performance, a plausible interpretation was that the selection of the particular strategy led to more efficient retention in the long-term memory stores, used in the case of the 2nd performance, happening after the 15-minute distraction period. This assumption was not validated by the same model applied to the 3rd performance of the piece, on the next day; it is important to note, however, that data from the 3rd performance were much fewer and thus were not appropriate for analysis.

The remaining predictors that were specific to the Atonal piece did not appear to have a significant effect in either of the first two performances; both the frequency of atonal music practice and the familiarity with atonal music yielded non-significant results. Instrument type, on the other hand, remained a significant predictor for performance, retaining the same ranking of instrument types as the Tonal piece: string and wind instrument (monophonic instrument) players appeared to obtain the highest scores, followed by pianists and then by singers.

At this point, after having established significant predictor variables for the 1st and the 2nd performance, the change in scores between the two performances was examined. Change in performance scores was calculated by subtracting the Total score of the 1st performance from the Total score of the 2nd performance. The resulting number, Δ_{ba}, could have either a positive or a negative value; negative values indicated that performance had deteriorated between the 1st and the 2nd time, whilst positive values indicated that performance had improved the 2nd time, resulting in a higher score. The difference in scores was also calculated for the 1st and 3rd performance, signified by Δ_{ca}.

MANOVA and multiple regression analyses did not reveal any significant predictors for the change in scores: amongst tried predictors were Instrument type, Memorisation frequency, Solfège use (implicit and explicit) and Memorisation difficulty. Certain interesting and potentially useful observations can be made, however, by comparing Δ_{ba}

and Δ_{ca} values between Solfège use groups, both implicit and explicit. Note the meaning of (I) and (E) distinctions in each group: the Solfège (I) group includes all participants who came from solfège countries; the Solfège (E) group includes participants who came from solfège countries and audibly used solfège during the 10-minute practice; the Non-solfège (I) group includes all participants from non-solfège countries; the Non-solfège (E) group includes all participants from non-solfège countries plus participants from solfège countries who did not use solfège audibly during the 10-minute practice session. Although such comparisons cannot replace statistically significant results of analysis of variance or multiple regression tests, they can nevertheless provide a useful basis for further testing, which will be more target-specific, controlled for independent variables and using a bigger sample. The total numbers of participants in each Solfège use group for each performance completed are shown in Table 3.22:

N	Tonal piece		Atonal piece	
	1st and 2nd performance	3rd performance	1st and 2nd performance	3rd performance
Solfège (I)	39	17	41	15
Non-solfège (I)	41	37	42	33
Solfège (E)	29	13	30	12
Non-solfège (E)	51	41	52	36
Total	80	54	83	48

Table 3.22: Number of participants who completed the respective sessions for the tonal and atonal pieces. AP participants are excluded from the above count.

Table 3.22 reveals that the total number of participants completing all three performances was considerably lower for both the Tonal and the Atonal pieces, compared to the number of participants attending the first experimental session, which included the 1st and the 2nd performance; especially solfège participants appear to have failed to complete the 3rd performance in their vast majority. The reason is that, apart

from country 2a (non-solfège, 30 participants), the rest of the experiments were organised to take place within a very strict timeframe: the result was that very few students were able to accommodate the next-day session into their schedules. As such, the participants in the 3rd performance came mainly from the 2a group for non-solfège participants, while for the solfège group there were fewer participants from each of the three countries. Taking these demographics into consideration, it is useful to examine Table 3.23, which gives the cumulative percentage of participants who scored positive Δ_{ba} and Δ_{ca} values for both pieces:

%	Tonal piece		Atonal piece	
	Δ_{ba}	Δ_{ca}	Δ_{ba}	Δ_{ca}
Solfège (I)	64.1%	64.7%	58.5%	40%
Non-solfège (I)	58.5%	43.2%	40.5%	30.3%
Solfège (E)	58.6%	61.5%	53.3%	33.3%
Non-solfège (E)	62.7%	46.3%	48.1%	33.3%
Total	61.2%	50%	49.4%	33.3%

Table 3.23: Percentage of participants who scored higher in the 2nd than in the 1st performance (Δ_{ba}) or in the 3rd compared to the 1st performance (Δ_{ca}), thus yielding positive Δ values.

The total percentages reveal that later performances of the Atonal piece deteriorated or stayed the same for most participants, whereas this was not the case for 2nd and 3rd performances of the Tonal piece. In fact, comparison of Δ_{ba} values, between the 1st and 2nd performance of the Tonal piece, shows that performance improved for most subjects, across groups; this might reflect the mediation of long-term memory mechanisms, which may have been activated during the 15-minute distraction period.

The biggest differences between solfège and non-solfège groups are observed in Δ_{ca} values: 64.7% of participants from solfège countries appear to have done better in the last performance of the Tonal piece than they had done in the 1st, while this percentage

was 43.2% for the non-solfège group. Although this difference is reduced in the comparison between solfège subjects who used solfège audibly during practice and all the other subjects, it still remains over 10 percentage units. These differences, of course, can receive numerous interpretations and, since they are not validated by statistical tests, causality is not tenable; they do provide, however, a basis for further research in the field, as the large differences between the groups could be justified by an advantage provided by solfège use in the long-term retention of the piece: given the restricted timeframe of the tasks, the use of solfège syllables may have helped the rapid formation of a more stable memory trace, which prevented performance deterioration.

Arguably the most interesting finding of Δ frequencies is that, for the majority of the subjects, the 2nd performance of the Tonal piece, after the 15-minute distraction period, was better than the 1st performance, which had happened immediately after memorisation. This could reflect the employment of long-term memory mechanisms in the 2nd performance and has potential pedagogical implications for music teachers and students who are interested in honing their memorisation skills and efficiency.

3.4. Discussion

Statistical analyses showed that, for both the Tonal and the Atonal piece, the most significant predictor of the memorised performance score was the type of instrument they played (Tables 3.16, 3.17, 3.18 and 3.19): wind and string players appeared to be the most competent memorisers, followed by pianists, with vocalists in the last place. This was accurately reflected in questionnaire responses regarding self-rating in memorisation skills (see Appendix 3): participants appeared to make realistic estimations of their memorisation competence, with only 31.3% of the singers assessing it as 'Good' or 'Extremely good', followed closely by pianists with 36.7%, whilst the equivalent percentages for string and wind players were 52.9% and 42.9%, respectively. The self-rating of participants' memorisation skills also appeared significantly correlated with their performance scores: τ values were all found significant at p $\leq$ .01 for the 1st and 2nd performances of both the Tonal and the Atonal piece.

Apart from the self-rating of memorisation skills, however, a brief examination of qualitative data from the questionnaires does not justify the singers' and pianists' low scores: although memorisation difficulty, as depicted in participants' responses, appeared significantly correlated with performance on the Tonal pieces and the 2nd performance of the Atonal piece at p < .05, the 1st performance of the Atonal piece did not have a significant correlation with Memorisation difficulty, which puts the consistency of the finding in question, at least for the Atonal pieces. Moreover, individual instrumental groups' responses in the Memorisation difficulty question reveal a relative contradiction with performance results: 46.9% of the singers and 43.3% of the pianists responded they found music memorisation in general 'Very difficult' or 'Difficult', while the equivalent responses for string and wind players were 58.8% and 50%, respectively. These findings, along with correlation tests, indicate that, although individual participants' attitudes towards memorisation were associated with how well they did on the task, at least for the Tonal piece, singers and piano players, as groups, did not find memorisation more difficult than wind and string players; this suggests that low performance scores were not likely to be a result of differences in cognitive abilities, but rather a consequence of technical, performance-related issues.

An example of a practical factor potentially affecting performance was time: questionnaire responses showed that the restricted timeframe within which participants were asked to memorise the 8-bar pieces had played an important role in different instrumental groups' performance scores. Figure 3.1 displays the percentage frequencies of participants who responded that they had either 'Plenty of time' or 'Just about enough time' in order to memorise the piece. These percentages are over 50% for both the string and the wind player groups, while less than 30% of the singers thought they had enough time to learn either the Tonal or the Atonal piece. The lowest frequency is displayed in the pianists' group, with only 3.4% (1 participant) responding they had enough time to memorise the pieces. These findings indicate a strong tendency amongst pianists and singers to take more time in order to memorise a given piece. The reason, however, that this variable was not entered in any factorial or regression analyses with performance scores was that it was based on the subjective opinion of the participants, as it had been depicted on the questionnaire that they had filled in *after* the task. Timings in the experiment were controlled variables, based on extensive piloting; the

fact that many participants, from all instrumental groups, managed to complete the task successfully and on time suggested that the questionnaire responses regarding time adequacy were, at least partly, a reflection of the participants' estimation of their performance and their time-management efficiency in the task. This was also confirmed by the 10-minute practice transcriptions, which revealed poor time-management in most cases of participants who had answered they would require 'a lot more time' to memorise the piece, compared to all other participants.

Participants who responded they had adequate time to memorise the piece
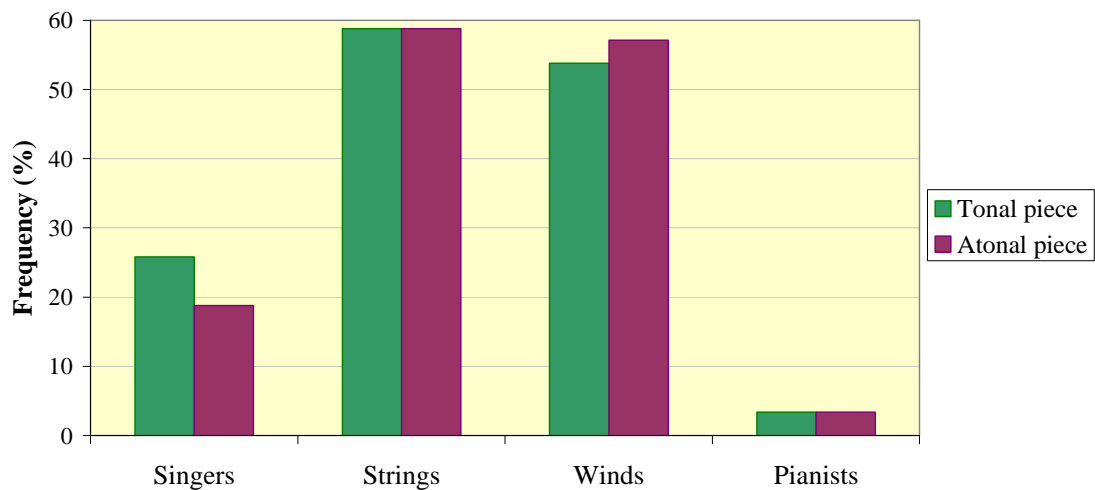
Figure 3.1: Instrumental group responses to the question: 'Did you feel you had enough time to memorise the piece?'. Only positive responses are recorded in the graph; the green bars represent percentage frequency for responses regarding the Tonal piece, whereas purple bars represent frequency of the positive responses for the Atonal piece.

Regression analysis showed Memorisation frequency to be a significant predictor of performance scores only for the 1st performance of the Tonal piece, while it was found non-significant for the 2nd performance of the Tonal as well as for both performances of the Atonal piece (Tables 3.16, 3.17, 3.18 and 3.19). This ambiguous finding for the tonal piece was cross-validated with questionnaire responses, categorised by instrumental group; percentage frequencies of participants who responded 'Always/Almost always' or 'Often' in the question regarding how often they memorise music for performance (Appendix 3) are shown in Figure 3.2. According to these

frequencies, singers appear to memorise music for performance more often than all the other groups; pianists are the group with the largest percentage of their population in the low-frequency group. This questionnaire finding suggests that, although memorisation frequency was an important factor in shaping performance scores, it was not sufficient in itself to overcome restrictions posed by the nature of the instrument performed.
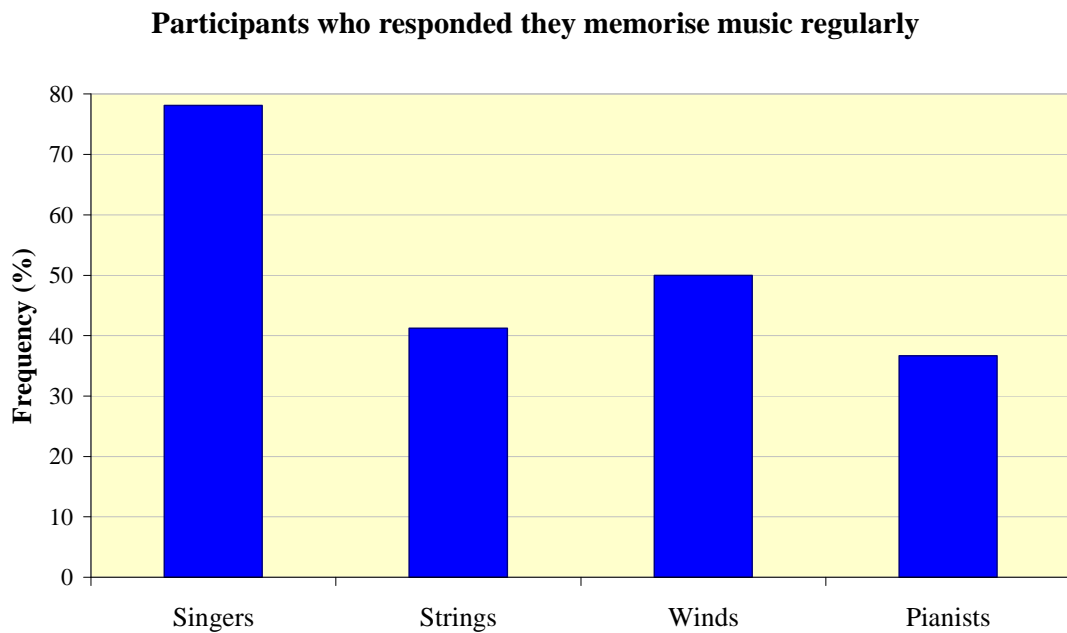
Participants who responded they memorise music regularly

Figure 3.2: Instrumental group responses to the question: 'How often do you memorise music for performance?'; possible responses included 'Always/Almost always', 'Often', 'Occasionally' and 'Never/Almost never'. The bars in the graph represent percentage of responses given for the first two options only.

Another significant predictor emerging from the multiple regression model for the Tonal piece was the use of a specific strategy (Table 3.20); participants who had responded they used a particular strategy for memorising music appeared to perform better than the ones who did not identify a strategy they used. There is, of course, a point of consideration regarding this predictor: as the responses were elicited *after* the memorisation task had taken place, it is possible that some participants may have rationalised their approach to the piece *in retrospect*, so that they gave an answer they felt was more appropriate both to the task and to their status as music students. Bearing this potential restriction in mind, it is interesting to look at how the 'analytical' strategy,

identified as the practice of the piece 'without the instrument, determining its structural elements' (see Appendix 3), was distributed across instrumental groups: 12.5% of singers stated they used this strategy, versus 17.6 % of the string players, 21.4% of the wind players and 23.3% of the piano players. The 'analytical' strategy was the only specific strategy that appeared to be a significant predictor in the multiple regression model for both performances of the Tonal piece. Although the percentage of singers using this strategy is the lowest amongst instrumental groups, differences are not very large; along with its significance in the regression model, however, this strategy appears to bear a relative advantage compared to other strategies. This is also consistent with findings from the music memorisation literature: Chaffin and Imreh (1997) and Ginsborg (2002) identify the memorisation based on the knowledge of the music's structural characteristics as the most reliable way of music memorisation, employed as the strategy of choice by most professional musicians and recommended by music teachers over the other types of music memorisation.

In the Atonal piece, the only specific memorisation strategy employed by participants, which also appeared to be a significant predictor of performance, was chunking, defined as 'dividing in sections and learning them separately' (Appendix 3); chunking appeared significant only for the 2nd performance of the Atonal piece (Table 3.21). As in the case of memorisation frequency, the questionnaire responses combined with practice transcriptions contradicted this result: the percentages of subjects who had actually used chunking by instrumental group were: 31.3% for singers, 22.6% for monophonic instrument players and 20% for singers. On the other hand, correlation results between the use of chunking and Δ_{ba} scores for the Atonal piece indicated that piano players who had used chunking had improved in the 2nd performance, $\tau = .40$, $p < .05$. The correlation of chunking and Δ_{ba} scores was found non-significant for the other instrumental groups; these results suggest that, although chunking may have provided a long-term benefit for the memorisation performance of those who used it, this effect was not strong enough to overshadow changes in performance triggered by performing a certain type of instrument.

The insignificance of factors such as the frequency of performing atonal music and the familiarity with atonal music (Table 3.21), along with the consistent significance of Instrument type, indicate that, as for the Tonal piece, there could be practical or

performance-related factors affecting performance scores. One such factor could be related to the experimental procedure itself and the fact that performances were audio recorded. Although there was a relevant question in the questionnaire, intending to measure the extent to which the recording affected participants' performance, the responses were not included in factor and regression analyses for the same reason as for responses regarding time adequacy: as the participants were asked after the task had taken place, it was possible that their response was affected by their assessment of their performance and therefore consisted an unreliable measure of how much recording had actually had an effect on them. Nevertheless, questionnaire responses regarding the effect of recording were examined: 77.2% of participants answered that the recording had not affected their performance in the tonal piece, while the respective figure for the atonal piece was at 83.1%. The extent to which recording affected the remaining participants' performance is shown in Figure 3.3:

To what extent did the fact you were being recorded affect your performance?
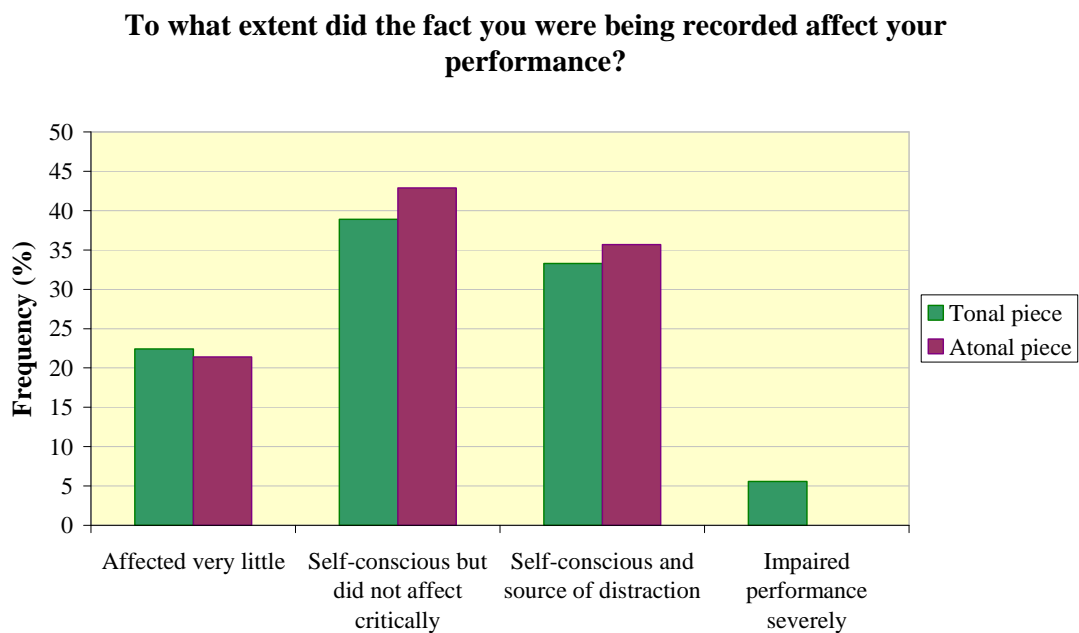
Figure 3.3: Participants could choose one of the following options (see Appendix 3): a) It affected me very little, b) It made me self-conscious but did not critically affect my performance, c) It made me self-conscious and was a source of distraction from the task, d) It impaired my performance severely, or e) Other.

Figure 3.3 shows that even the participants who found that recording had affected their performance, believed in most cases that this effect was relatively small: out of 93

participants, only 7 students in the Tonal and 5 in the Atonal piece stated that recording had critically affected their performance.

Overall, the vast majority of the participants appeared to value memorisation skills, with 41.9% of them rating them as 'Extremely important', 51.6% as 'Important' and 4.3% as 'Neither important nor unimportant'; only 2 participants answered they considered memorisation skills to be unimportant for a musician and no participants selected the 'Extremely unimportant' option. The distribution of answers for each instrumental group is shown in Figure 3.4: singers and wind instrument players appear to be the groups which attribute the highest importance to memorisation skills, while, on the other hand, there were some, albeit very few, pianists who considered memorisation skills unimportant.
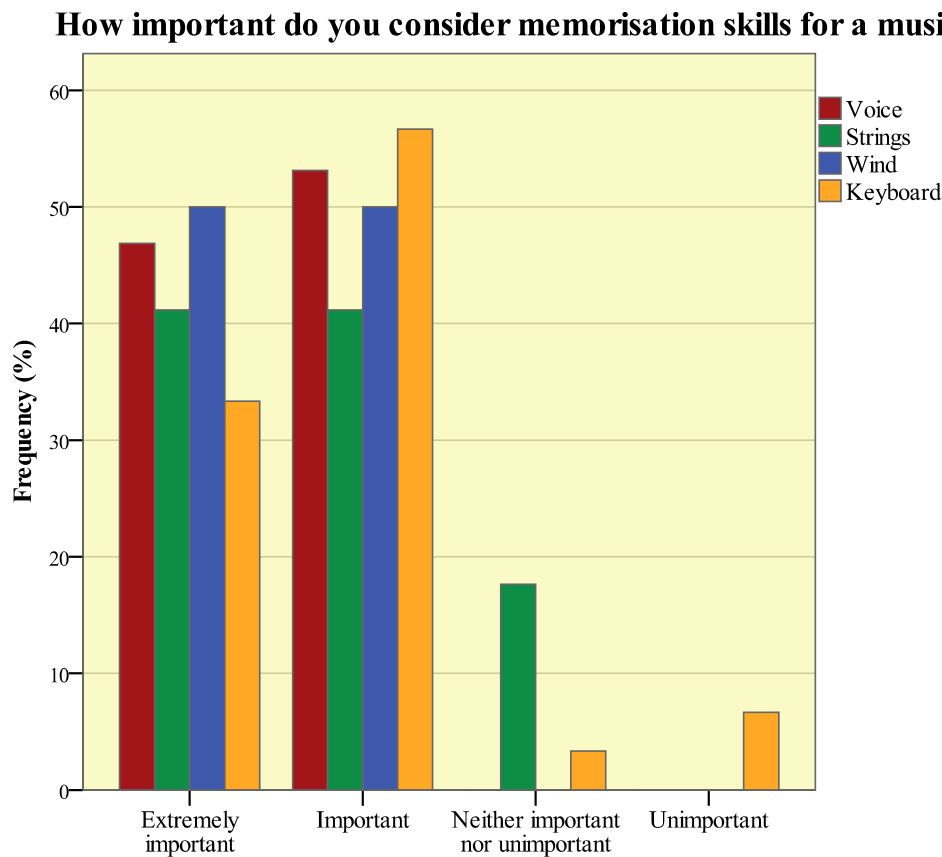
Figure 3.4: Participants' responses regarding the importance of memorisation skills, categorised by instrument. Note than the 'Extremely unimportant' option was not selected by any participants.

Despite the fact, however, that most participants appeared to consider memorisation skills vital for musicians, most of them appeared to have cultivated these skills on their own: 75 out of the 93 participants answered that they had never been taught how to memorise music. The self-taught approach to memorisation could also provide an explanation for the inconsistency between questionnaire responses and the recorded reality regarding memorisation strategies used during practice: 37.8% of participants failed to use the strategy they had described in the questionnaire for the Tonal piece and 27.5% of participants failed to use it during their practice for the Atonal piece. Bearing in mind that 84.5% of participants declared they made use of a particular strategy in memorising music and most of them selected chunking as their strategy of choice, it appears that most music students seem to know the 'correct' way towards efficient memorisation in theory and perhaps guided by intuition; the lack of a systematic approach, however, starting with the absence of memorisation-specific teaching by their instrumental and vocal instructors, results in flaws in its practical application.

The first hypothesis tested in the experiment stated that:

1. **Musicians who know and use solfège will perform better in a music memorisation task than musicians who do not know and do not use solfège.**

This hypothesis was not supported by analysis findings; the distribution of memorisation scores for the solfège and the non-solfège groups appeared similar, unaffected by either the implicit or the explicit use of solfège.

The second hypothesis of the experiment stated that:

2. **The positive effect of solfège utilisation in the memorisation process will be highest amongst singers and string players and lowest in piano players.**

As the previous hypothesis failed to be validated, this hypothesis was not supported by analysis findings, either; no positive effect of solfège use, implicit or explicit, was detected in any of the instrumental groups. The instrument performed, however, appeared to play an important role per se in memorisation performance: wind and string players yielded significantly higher results than pianists and singers.

The third hypothesis tested in Experiment 1 stated that:

3. The positive effect of solfège utilisation will be stronger for atonal music than for tonal music.

This hypothesis was not supported, either; again, instrumental group appeared to be the most significant factor affecting performance scores in the memorisation of the Atonal piece.

Although the comparative benefits of solfège use in music memorisation were not supported by the analyses, an important note was that a close examination of the 10-minute practice transcriptions revealed that 25.5% of solfège participants did not really use solfège in the Tonal piece and 23.9% did not use it in the Atonal piece. Transcriptions, of course, could only provide information regarding explicit, or audible, solfège use: the precise role of solfège note names during the process of familiarising oneself with a new piece cannot be investigated thoroughly with the behavioural/ observation methods used in the present study. In most cases, however, observations regarding solfège use could be quite conclusive: when, for example, non-singers could be heard playing the piece incessantly during the 10-minute practice, it was considered impossible to have consciously and purposely used solfège towards memorisation. There were still cases, though, in which solfège use and the strategy employed in general was extremely hard to decipher from recordings. This led to the conclusion that the real-life approach adopted in the methodology of this study (Chapter 2.0), despite its advantages, had resulted in the considerable augmentation of unsystematic variation within the experiment, which, in turn, affected analysis and results. This effect will be discussed further in Chapter 6.

Despite the fact that not all solfège participants had actually used solfège explicitly during practice, only one participant in the Tonal and one in the Atonal piece stated that in the questionnaire; all the other participants responded to questions regarding solfège use in the task as if they had used it. This can be explained either by assuming that some of the participants were referring to the effects of implicit solfège knowledge and use on the task, or by interpreting some of the responses as an attempt to conceal the fact that they had not, in fact, used solfège. Questionnaire responses regarding the effect of solfège use in the task by solfège participants are summarised in Figure 3.5:

How did the use of solfege affect your memorisation?
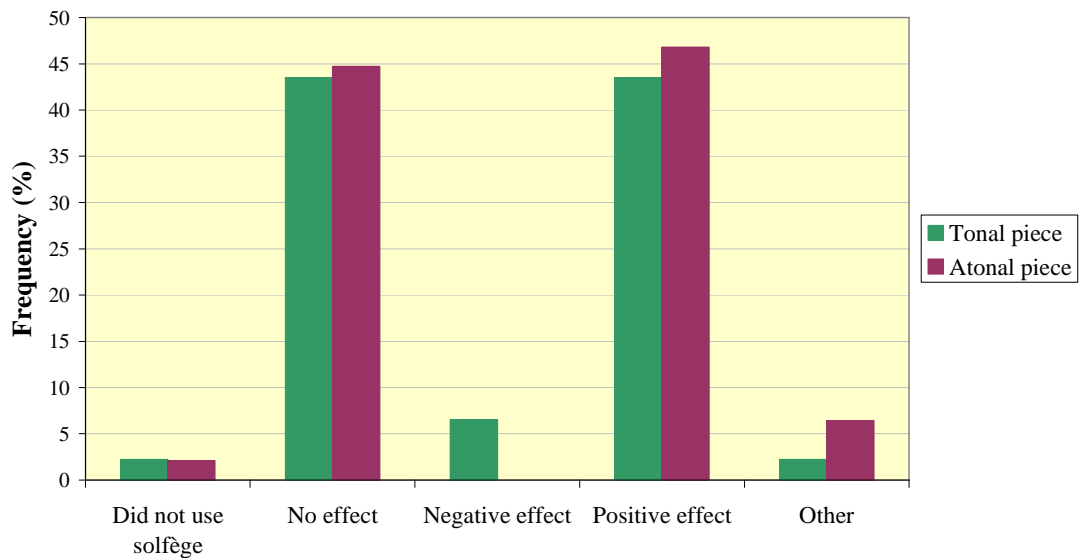

Figure 3.5: Percentage frequencies of participant responses regarding the effect of solfège on their memorisation of the Tonal and the Atonal piece.

Figure 3.5 shows that most participants, for both pieces, believed that solfège had either no or a positive effect on their memorisation. Out of the 20 subjects who had selected the 'No effect' option, 9 had not used solfège explicitly and the remaining 11 had used it. Possible interpretations for the 'No effect' answers included the following:

1. Subjects who had not used solfège explicitly may have been implying their lack of solfège use by selecting the 'No effect' option.

2. 45.5% of the subjects who had used solfège explicitly and had found no effect on their memorisation were also placed in the high-frequency group in respect to their use of solfège for memorising music, according to their questionnaire responses. Since these participants were accustomed to using solfège for memorising, the 'No effect' response may have meant that, since they followed their regular mode of memorisation, they found no changes in the process or the result.

3. Some of the 'No effect' responses may be explained by considering data regarding the self-taught approach most participants had on music memorisation: participants who had not been taught how to memorise music and were not directly aware of a particular strategy they were using, were expected

to be less aware of the factors affecting their memorisation performance than their counterparts.

Moving on to the 'Negative effect' responses for the Tonal piece, it is important to note that all three subjects who gave this response were singers; in a subsequent question, where they were asked to elaborate on their response, all three of them explained that, because they thought they had to use solfège in their performance after 10 minutes, it made them pay more attention to note names rather than the music itself. This had happened despite the fact that written and oral instructions stated explicitly that participants did not need to use solfège when performing the piece from memory for the researcher. The small number of participants giving a 'Negative effect' response, the fact that such responses were elicited only for the Tonal piece and the apparent amount of unsystematic variation, established in previous analyses, led to the conclusion that individual differences that had not been measured in the experiment had most likely played a crucial role in shaping participants' performances as well as responses.

An important assumption of the original hypotheses tested in the experiment was that solfège participants would be comfortable in using solfège labels in order to memorise the pieces. This belief was proved to be false for many of the subjects: according to questionnaire responses, 39% of the solfège participants were placed in the low-frequency group regarding solfège use in everyday music practice and 43.9% were in the low-frequency group regarding solfège use in music memorisation. This was confirmed by the interaction with subjects prior to the experimental procedure: many of the subjects required clarifications as to how or where they were required to use solfège, showing an obvious discomfort at the prospect. Although their reaction may have been partly caused by anxiety regarding the task itself, it was apparent that solfège is not as automatic a procedure as presumed, for all students. This was confirmed by the relevant questionnaire responses, shown in Figure 3.6:
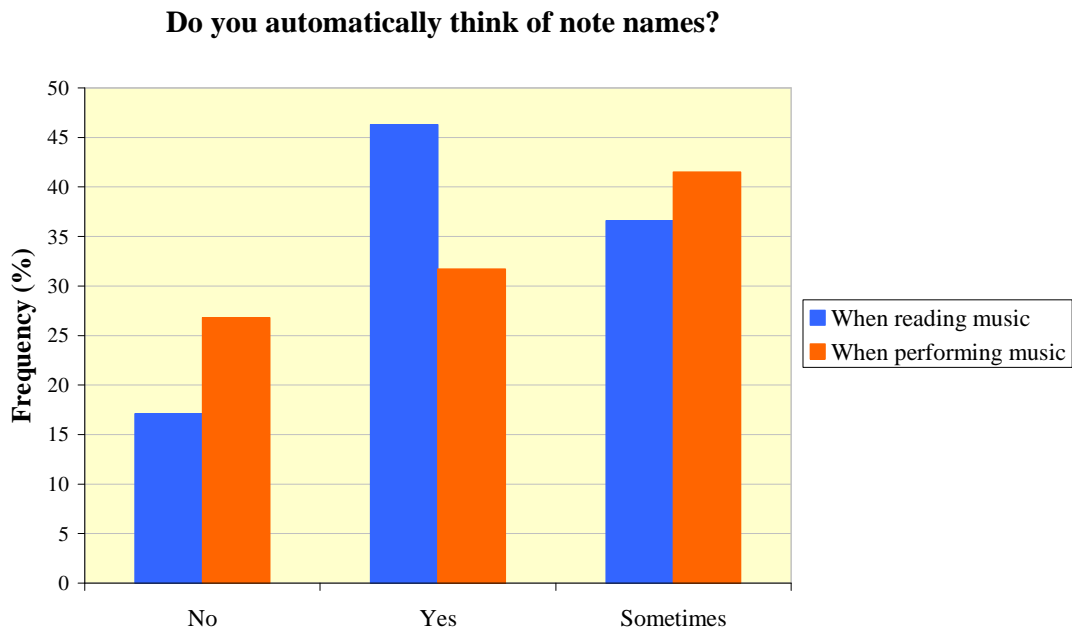
Do you automatically think of note names?

Figure 3.6: Participant responses to the questions: 'Do you automatically think of note names when reading music?' (blue bars) and 'Do you automatically think of note names when performing music?' (orange bars). This question was directed to solfège participants only, therefore 'note names' denotes only solfège syllables.

17.1% of participants answered they do not think of solfège note names when reading music and 26.8% that they do not automatically think of them when performing music. It is important to stress again at this point the obvious inadequacy of the method used to assess the automatic use of solfège labels, as participant responses reflect what they *believe* happens, which may not always be the same as what *actually* happens in their brain when they encounter a musical score. Bearing this restriction in mind, the participants' responses reveal that solfège note names are not automatic during music reading and performance for all musicians; following analysis results regarding factors affecting memorisation, a tentative conclusion can be drawn that solfège, just like any other skill, needs to be cultivated in order to become useful and productive as a method of approaching music. Its unique nature as a referent, as described in Chapter 1.2, would be best investigated using alternative experimental methods, such as brain imaging or behavioural testing like the one described in Chapter 5. Music memorisation, on the other hand, as examined in the present experiment, proved to be an extremely complex, multi-factorial procedure, so that solfège in itself did not appear to induce traceable

changes to the results; an alternative approach to investigating the hypothesised comparative advantage of solfège in real music memorisation circumstances will be presented in the next chapter, within the framework of a repeated-measures design.

Chapter 4

Experiment 2

The effect of solfège instruction on music memorisation: A longitudinal study

4.0. Aims and hypotheses

Experiment 2 was designed to provide a repeated-measures perspective to the effect of solfège use on memorisation tasks: while Experiment 1 had been designed to test absolute differences between solfège and non-solfège musicians, whereas Experiment 2 aimed to test relative differences in the performance of the same musicians, before and after they had been taught solfège. The general aims of the experiment were:

1. To examine the effects of solfège use on a music memorisation task performed by adult musicians who have been taught solfège recently.

2. To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians who have been taught solfège recently.

In order to achieve these aims, Experiment 2 tested the following hypotheses:

1. **Musicians who learn solfège as adults are able to use it effectively in order to memorise music.**

   As explained in Chapter 2.1, solfège, and especially fixed-do solfège, is normally taught during childhood, over an extended period of time of a minimum 5 years. The solfège lessons that took place for the purposes of Experiment 2 proved successful in familiarising students with solfège syllables and honing their aural skills (see section 2.1.3); the main objective of the experiment, however, was to check whether or not solfège could be used at request for the completion of specific memorisation tasks.

2. **Learning and using solfège will improve performance in a music memorisation task compared to not using solfège in the same task.**
   The use of solfège syllables was expected to offer an additional encoding for the musical stimulus and therefore an advantage in the execution of the memorisation task.

3. **The positive effect of solfège utilisation in the memorisation process will be highest amongst singers and string players and lowest in piano players.**
   As in Experiment 1, the use of solfège was expected to be more challenging for pianists, because of the existence of more than one melodic line which they have

to sing, and wind instrument players, because of the physical requirements of their instruments, which prevent the simultaneous execution of performance and solfège singing.

4. **The positive effect of solfège utilisation will be stronger for atonal music than for tonal music.**

   Because of the presumed familiarity of most subjects with tonal music, as opposed to atonal music, atonal stimuli were expected to present more of a challenge in their memorisation within a restricted timeframe. For this reason, the proposed additional encoding offered by solfège was expected to create an advantage in the abstraction and the construction of mnemonic patterns for atonal stimuli.

4.1 Method

4.1.1. Design

The study employed a mixed design, with three independent variables:

1. **Solfège use**, with two levels: solfège and non-solfège.
2. **Instrument performed**, with four levels: voice, string, wind, piano.
3. **Type of piece performed**, with two levels: tonal and atonal.

The above were the independent variables which were controlled in the experiment: participants were selected and categorised according to these three factors. More independent variables were identified in the process of data analysis from the questionnaires; the added variables were also used in the main analysis for the experiment.

The within-subjects variable was the final score on the memorised task, with two levels: performance in Phase 1 and performance in Phase 3 (see section 4.1.4). All participants were asked to perform the task twice, one with a tonal and one with an atonal piece; the

order in which they received the pieces was random, controlling for repeats of material between back-to-back experiments (see section 2.2.3.3). This controlled randomisation was employed so that subjects would not overhear the previous participant performing the piece they were about to memorise or recall: in order to achieve this, a list of the 9 possible combinations of the 3 tonal and the 3 atonal pieces was used; when an overlap in material occurred between participants with back-to-back sessions, the second participant received the piece dictated by the next item on the list.

4.1.2. Participants

A total of 30 university music students, 19 females and 11 males, participated in the study: 18 students formed the experimental group and the remaining 12 students were in the control group. The control group included 3 singers, 3 string players, 1 wind instrument player and 5 piano players. The experimental group included 7 singers, 4 string players, 4 wind instrument players and 3 piano players; 3 of the singers, one string player and one wind player from the experimental group failed to complete the experimental manipulation and were subsequently moved to the control group. Furthermore, one singer from the experimental group failed to complete the experiment and was subsequently eliminated from the analyses.

All participants were full-time undergraduate students in the Department of Drama and Music in the University of Hull at the time: 28 of them were English and 2 were international students. Participants were all aged between 18 and 23 years old, with a mean age of 19.8 years. All 30 students volunteered for the experiment and signed the relevant consent forms; in addition, the 12 students who participated successfully in the experimental group were offered module assignment credit.

Data from all 30 participants in Phase 1 (see section 4.1.4) were also used in the between-groups design described in Chapter 3.

4.1.3. Materials

The materials used in this study were the same as the ones described in Experiment 1 (section 3.1.3): musical material consisted of 6 pieces: three tonal pieces (T1, T2, T3) and three Atonal pieces (A1, A2, A3).

Music memorisation tasks were run in a university practice room, equipped with an acoustic upright Yamaha piano (model U3). All participants were recorded using an Olympus VN-6500 PC digital voice recorder with an additional Olympus ME15 electret condenser microphone. Solfège instruction took place in a seminar room equipped with an upright piano and a whiteboard (see section 2.1.2.2); each student was given a copy of the solfège handbook (Appendix 1), put together specifically for the purposes of the study.

All participants were asked to complete a questionnaire for each stimulus they received in the memorisation tasks (see Appendix 3). The experimental group was also asked to complete an anonymous feedback form, regarding their views on solfège training, as well as a questionnaire regarding solfège lessons (Appendix 1).

4.1.4. Procedure

An overview of the experimental design is provided below:

Phase 1 (October 2009)

    A. i) Memorisation and 1st performance

        Learn T1 (10 minutes) – Perform T1 from memory.

    ii) Distraction (immediately after 1st performance)

        Complete questionnaire for T1 and perform distraction activities (15 minutes).

    B. Recall practice and 2nd performance (immediately after Distraction)

        Practise T1 without score (2 minutes) – Perform T1 from memory.

    C. Recall practice and 3rd performance (24 hours later)

        Practise T1 without score (2 minutes) – Perform T1 from memory.

Phase 2 (Experimental group only – October 2009 to May 2010)

Intensive solfège course led by the researcher (see Chapter 2.1.2).

Phase 3 (May 2010)

    A. i) Memorisation and 1st performance

        Learn T2 (10 minutes) – Perform T2 from memory.

    ii) Distraction (immediately after 1st performance)

        Complete questionnaire for T2 and perform distraction activities (15 minutes).

    B. Recall practice and 2nd performance (immediately after Distraction)

        Practise T2 without score (2 minutes) – Perform T2 from memory.

    C. Recall practice and 3rd performance (24 hours later)

        Practise T2 without score (2 minutes) – Perform T2 from memory.

Phases 1 and 3 were identical in structure and were repeated once each for each participant, using atonal pieces (A1, A2 or A3). T1 and T2 are used in the above descriptions as examples; the subjects in the study were given any combination of two different pieces between T1, T2 and T3 (or A1, A2, A3 for the atonal stimuli). For a detailed account of each separate step in Phases 1 and 3, see section 3.1.4.

4.1.5. Transcriptions and scoring

All twelve performances of each participant (three memorised performances for the tonal and three for the atonal piece from Phase 1 and the same amount from Phase 3) were transcribed from the voice recorder to manuscript paper[1] and then marked according to pitch accuracy, rhythm accuracy, structural accuracy and fluency. The pitch and rhythm scores were added up and created an overall 'technical accuracy score' for the piece; subsequently, all three scores, technical, structure and fluency, were converted into percentage scores and the resulting scores were added up using two different algorithms. For a detailed description of the scoring system see section 2.3.3.

---

[1] The decision to produce hand-written transcriptions rather than using the relevant software is explained in section 2.2.1.5.

Each participant was recorded in two 10-minute private practice sessions – one for the tonal and one for the atonal piece – in Phases 1 and 3, producing a total of four 10-minute session recordings for each participant: these were transcribed into text, as well as the two 2-minute recall sessions for each piece in Phases 1 and 3.[2] For the present study, the only finding used from the transcriptions of the practice sessions was the comparison between the strategies the subjects had actually used to memorise the music during the 10-minute practice session and the strategy they had stated they used in the questionnaire.

4.2. Preliminary analysis: piece comparability

Although the tonal pieces (T1, T2, T3) and the atonal pieces (A1, A2, A3) were compared in Chapter 3.1., for Experiment 1, and were found equivalent in terms of the memorised performance scores they produced, the three pieces in each group were compared again for the purpose of the repeated-measures design, in order to establish that the particular participants' performances were not significantly affected by the specific piece performed.

The distributions of performance scores for all pieces, across performances, showed significant deviations from normality, consistent with findings from the study using the same pieces with a larger sample (Chapter 3.1). For this reason, non-parametric tests were used again, in order to compare the distribution of scores between different pieces in the tonal and the atonal condition.

For the tonal pieces, the Kruskal-Wallis test statistic was found non-significant at $p > .05$ for the 1st, 2nd and 3rd performances of the tonal piece in Phase 1 and Phase 3; this suggested that the specific piece performed did not have a significant effect on the memorised performance. The same was true for all performances of the atonal pieces: the Kruskal-Wallis test statistic was, once more, found non-significant ($p > .05$) across

---

[2] A sample transcription of a practice session is provided in section 3.1.5.

performances, in both Phases, suggesting that A1, A2 and A3 produced similar memorised performances.

Questionnaire data corroborated the hypothesis that pieces within the tonal and the atonal group were equal in terms of both technical and memorisation difficulty; response summaries are given in Table 4.1 and Table 4.2:

Participant responses regarding the technical difficulty of the piece

Response / Piece tested	Very difficult / Difficult	Neither difficult nor easy	Very easy / Easy
T1	2	5	14
T2	1	7	12
T3	3	7	8
A1	4	4	12
A2	5	5	10
A3	5	7	7

Table 4.1: Participant responses regarding the tonal and the atonal pieces' technical difficulty. Subjects choosing the 'Very difficult' or the 'Difficult' option were pooled together; the same was done for subjects choosing the 'Very easy' or 'Easy' options. The reason for this merge was that these responses were considered to represent different levels of the same category: the former group included all subjects who had found the piece challenging and the latter included subjects for whom the piece did not present particular challenge. For tonal pieces, N=21 N=20 and N=18 for T1, T2 and T3 respectively. For atonal pieces, N= 20 for A1, A2 and N=19 for A3.

Participant responses regarding the memorisation difficulty of the piece

Response Piece tested	Very difficult / Difficult	Neither difficult nor easy	Very easy / Easy
T1	10	7	4
T2	9	7	4
T3	6	7	5
A1	13	4	3
A2	12	6	2
A3	8	7	4

Table 4.2: Participant responses regarding the tonal pieces' memorisation difficulty. Sample sizes for each piece are equal to sample sizes in the response regarding technical difficulty (Table 4.1); categorisation of responses is explained in Table 4.1.

Questionnaire responses regarding the technical and the memorisation difficulty of the pieces were very similar to the ones acquired in Experiment 1 (Tables 3.2 and 3.3). The relationship between responses regarding memorisation difficulty and piece was also investigated for the responses acquired in Phase 3 and was found non-significant at $p > .05$ both for the tonal, $L\chi^2 (6) = 5.92$, *ns*, and for the atonal pieces, $L\chi^2 (6) = 4.26$, *ns*.

Summary

Statistical analyses of performance scores and questionnaire responses showed that the variation between T1, T2, T3 and A1, A2, A3 pieces had not produced any significant effect on memorised performance scores. This was also in line with findings from Experiment 1 (Chapter 3.1), suggesting that performances in the three tonal and the three atonal pieces could be pooled together, respectively, in a single group and analysed as such.

4.3. Main analysis

The major objective of data analysis was to determine whether the use of solfège had an effect on memorised performance scores and, if that was the case, if the magnitude of the effect varied according to the instrument performed. In order to investigate this, the following three parameters were examined:

1. **Instrument type.** The participants were divided in four different categories: *i) vocalists*, including male and female first-study singers, *ii) string players*, including violin, viola, violoncello and double bass players, *iii) wind players*, including both woodwind and brass instrument players and *iv) pianists*. Taking the results of the between-groups design into consideration (see Chapter 3.3), according to which string and wind players produced similar patterns of results, these two categories were, on occasion, pooled together, so that the resulting categories were three: *i) vocalists*, *ii) monophonic instrument players*, including all string, woodwind and brass instrument players and *iii) pianists*.

2. **Solfège use.** As in the between-groups design (Chapter 3.3), this was examined in two levels: in the first level, subjects were divided into solfège and non-solfège groups according to whether they belonged in the experimental or the control group, respectively. In the second level, participants in the solfège group were filtered further according to the data from the 10-minute practice transcriptions: participants who had *actually* used solfège were separated from those who had not used solfège in an audible way for at least 1 minute during practice. As in Experiment 1, this categorisation was useful in distinguishing between the lower, or first, level of acquaintance with solfège note names, which does not necessarily imply use of these elements in musical reality, and the higher, or second, level of extended and cultivated familiarity with solfège labels, which is presupposed for the use of solfège in music practice. The distinction between the two levels may also be perceived as a potential distinction between the *implicit* and the *explicit* effects of solfège knowledge. For this reason, the labels *Solfège use (I)* and *Solfège use (E)* are going to be used hereon to denote whether the first-level (implicit) or the second-level (explicit) solfège knowledge is being investigated.

In the course of the analysis, the effect of other variables was examined as well, in order to clarify the main factors affecting performance; such factors were AP possession, memorisation frequency for the Tonal piece, and the level of familiarity with atonal music for the Atonal piece. Statistical analyses were conducted using three-way mixed ANOVAs, repeated contrasts and multiple regression analysis.

4.3.1. Repeated measures analysis

In line with the findings from the analysis of data from the 1st experiment (see Chapters 3.1 and 3.2), the distributions for performance scores within groups were found significantly non-normal in most occasions. F-statistics are known to be robust to deviations from normality only when group sizes were equal (Field, 2009); in this case, the sample sizes for groups categorised by Solfège use (I) were N = 17 and N = 12 for the control and the experimental groups, respectively. The 17 subjects of the control group included two AP participants, whereas there were no AP possessors in the experimental group; supported by findings from Experiment 1 (Chapter 3.3), indicating that AP possession affected performance results, and following the need to balance sample sizes, the 2 AP participants were removed from the analyses. Even after this correction, however, group sizes were still different, both in themselves and according to instrumental type: following the removal of AP subjects, the control group included 6 singers, 4 monophonic instrument players and 5 pianists, whereas in the experimental group there were 3 singers, 6 monophonic instrument players and 3 pianists. Bearing these restrictions in mind, ANOVA results were used with caution regarding their possible interpretations.

F-statistics for all three performances of both the Tonal and the Atonal piece in Phases 1 and 3, did not reveal a significant effect of solfège use (I) on performance scores, apart from one case: in the 3rd performance of the Tonal piece, there appeared to be a significant interaction between performance scores in Phase 1 and 3 and whether the participant had or had not learnt solfège, $F(1, 20) = 7.37$, $p < .05$; this interaction was not accompanied by a significant main effect of solfège use (I), $F(1, 20) = 0.43$, *ns*. It is important to note at this point that, apart from the violation of the normality assumption,

homogeneity of variance had also been violated; as, however, the non-solfège group, which was larger, had produced larger variances than the solfège group, the resulting F statistic was believed to be conservative (Field, 2009), and therefore reliable. Nevertheless, because of the violation of the normality assumption, this finding was only used as an incentive to investigate further the changes in scores in the 3rd performance in further analysis and was not interpreted per se.

Instrument category (4 groups), on the other hand, appeared to have a significant effect at p < .05, in all performances but the 3rd performance of the Atonal piece.[3] Repeated contrasts revealed that, as in Experiment 1, the string and wind group produced significant contrast interactions with both the singer and the pianist group; differences between singers and pianists were found non-significant in all cases.

4.3.2. Δ_Δ scores

In order to proceed with further analysis without increasing the risk of Type I and Type II errors, which would be a consequence of using mixed ANOVAs when assumptions had been violated, analysis was done using Δ scores (see Chapter 3.3) instead of actual scores. Δ scores represented change in performance scores calculated by subtracting the Total score of one performance from the Total score of another performance for the same piece, in the same Phase. For example, Δ_{T1ba} was calculated by subtracting the Total score of the 1st performance of the Tonal piece in Phase 1 from the Total score of the 2nd performance of the same piece, 15 minutes later. The resulting score could have either a positive or a negative value; negative values indicated that performance had deteriorated between the 1st and the 2nd performance, whilst positive values indicated that performance had improved, resulting in a higher score the 2nd time.

---

[3] All effects are reported as significant at p < .05. For the 1st performance of the Tonal piece there was a significant main effect of instrument type, $F(3, 21) = 4.83$, for the 2nd performance of the Tonal piece there was also a significant main effect of instrument type, $F(3, 21) = 3.72$, as well as for the 3rd performance of the Tonal piece, $F(3, 20) = 7.37$. For the 1st performance of the Atonal piece there was a significant main effect of instrument type, $F(3, 21) = 5.86$, for the 2nd performance of the Atonal piece there was also a significant main effect of instrument type, $F(3, 21) = 5.55$, but no significant effect was found for the 3rd performance of the Atonal piece, $F(3, 17) = 2.64$, *ns*.

These scores were calculated for the change between all pairs of performances of both pieces, in Phase 1 and in Phase 3. Subsequently, Δ scores for Phase 1 were subtracted from their counterparts in Phase 3, yielding a Δ_Δ score, which showed how much the subject's performance had changed over the Phases. The function of Δ_Δ scores can be made clearer by using an example, using the hypothetical scores provided in Table 4.3:

Phase 1			Phase 3		
T1a	**T1b**	**T1c**	**T2a**	**T2b**	**T2c**
600	660	630	700	660	680

Table 4.3: Scores assigned to a hypothetical participant, who completed all 3 performances for Phase 1 and Phase 3. T1 and T2 represent the two different tonal pieces the participant was tested on, while a, b and c signify the 1st, 2nd and 3rd performance of the same piece, respectively.

According to the scores in Table 4.3, Δ scores for this imaginary subject would be the following:

$\Delta_{T1ba} = 660 - 600 = 60$

$\Delta_{T1cb} = 630 - 660 = -30$

$\Delta_{T1ca} = 630 - 600 = 30$

$\Delta_{T2ba} = 660 - 700 = -40$

$\Delta_{T2cb} = 680 - 660 = 20$

$\Delta_{T2ca} = 680 - 700 = -20$

The above numbers reflect that the performance of this subject improved by 60 points in the second performance and then deteriorated by 30 points in Phase 1, while in Phase 3 the performance deteriorated the second time and then got better the last time.

Δ_Δ scores for the same participant would be:

$\Delta_{\Delta Tba} = -40 - 60 = -100$

$\Delta_{\Delta Tca} = -20 - 30 = -50$

$\Delta_{\Delta Tcb} = -20 - (-30) = 10$

As such, Δ_Δ scores can function as an overall assessment of the experimental manipulation that took place between Phases 1 and 3, indicating the nature and the extent of the effect it had had on the participant's performance, regardless of the absolute values. For example, the fact that in Phase 3 the 2nd performance deteriorated, while in Phase 1 it had improved, is reflected on the large negative value of $\Delta_{\Delta Tba}$. On the other hand, the fact that the 3rd performance showed less deterioration than the 2nd in Phase 3, compared to Phase 1, is reflected in the positive value of $\Delta_{\Delta Tcb}$: although in both cases performance deteriorated, the fact that it deteriorated less in Phase 3 is a relative improvement.

Δ_Δ scores can be, thus, viewed as providing 'standardised' values of score changes, independent of absolute values. These are also useful for diminishing possible effects between the different tonal and atonal pieces performed (T1, T2, T3 and A1, A2, A3 respectively): differences in technical difficulty, which could result in lower overall scores, are neutralised in Δ scores.

In order to make interpretation easier, Δ_Δ scores will be hereon named like simple Δ scores:

1. Δ_{Tba} , Δ_{Tcb}, Δ_{Tca} , will represent the change in scores between the 1st – 2nd, 2nd – 3rd and 1st – 3rd performance of the Tonal piece, respectively, between Phase 1 and Phase 3.

2. Δ_{Aba} , Δ_{Acb}, Δ_{Aca} , will represent the change in scores for between the 1st – 2nd, 2nd – 3rd and 1st – 3rd performance of the Atonal piece, respectively, between Phase 1 and Phase 3.

Positive values of these scores indicate that the score was higher in the latest performance, thus performance had improved; negative values, on the other hand, indicate that the score was lower in the latest performance, thus performance had deteriorated.

4.3.3. Analysis of Δ_Δ scores

Multiple regression analysis was conducted using each one of the Δ_Δ scores as the outcome variable, with two predictors: instrument type (3 groups) and solfège use (I). Residuals plots raised concerns regarding heteroscedasticity in the data, so bootstrapped samples were used. Δ_{Tba} scores were the only ones that yielded non-significant R^2 changes ($p > .05$); multiple regression statistics for Δ_{Tca} and Δ_{Tcb} are provided in Table 4.4:

Δ_Δ scores for the Tonal piece	Δ_{Tca}			Δ_{Tcb}		
	B	SE B[a]	β	B	SE B[a]	β
Step 1						
Constant	- 51.33	55.30		120.59	32.25	
Singers vs. Monophonic	- 159.39	118.30	- .29*	- 289.59	80.23	- .67**
Singers vs. Pianists	13.75	115.05	.02*	- 161.99	61.86	- .36***
Step 2						
Constant	- 124.58	80.37		65.85	51.25	
Singers vs. Monophonic	- 232.64	129.83	- .42**	- 344.33	79.98	- .80***
Singers vs. Pianists	4.60	127.92	.01*	- 168.83	71.17	- .38**
Solfège use (I)	219.75	105.31	.42**	164.22	64.14	.40**

Δ_{Tca}: $R^2 = .09$ for Step 1, $\Delta R^2 = .16$, $p < .05$ for Step 2.

Δ_{Tcb}: $R^2 = .35$ for Step 1, $\Delta R^2 = .14$, $p < .05$ for Step 2.

* $p > .05$, *ns*
** $p < .05$, *s*
*** $p \leq .01$, *s*
a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 4.4: Multiple regression statistics for Δ_{Tca} and Δ_{Tcb}. Instrument use (3 groups: singers, monophonic instrument players, pianists) and Solfège use (I) were used as predictor variables.

The values and significance of ΔR^2 scores reveal that the inclusion of instrumental type and solfège use improved the model significantly. More specifically, negative B values indicate that singers showed the most relative improvement between Phases 1 and 3, both between the second and third and between the first and third performance. It is important to remember at this point that the term 'relative improvement' does not necessarily imply that singers' performances overall got better between the 1st and the 3rd or the 2nd and the 3rd performance; rather, it suggests that, even in cases where performance deteriorated between performances, it tended to deteriorate less in Phase 3 than it had done in Phase 1. The difference between singers and pianists in relative improvement between the 1st and the 3rd performance was found non-significant; it was found significant, however, for the change between the 2nd and 3rd performance, with pianists more likely to show deterioration in performance scores than singers.

Implicit solfège use also appeared significant for improvement between performances in Phases 1 and 3; B values suggest that participants who had completed successfully the solfège course tended to show more relative improvement for the 3rd performance in Phase 3. It is important to note that the same model was tried with the substitution of implicit solfège use with explicit, for all Δ_Δ scores, but was found non-significant in all cases. This apparently contradictory finding will be examined further in the subsequent section, analysing Δ_Δ frequencies.

The same predictors, Instrument type (3 groups) and Solfège use (I), were entered in multiple regression models for Δ_Δ scores for the Atonal piece performances. For the Δ_{Acb} and Δ_{Aca} models, ΔR^2 values were found non-significant just over .05 (p = .052 for Δ_{Acb} and p = .054 for Δ_{Aca}); parallel to Δ_{Tba} , ΔR^2 for Δ_{Aba} was highly non-significant at p = .39 .

When, however, familiarity with atonal music was entered as a sole predictor in the analysis for Δ_{Aba}, the model reached significance. The categorisation of participants regarding their familiarity with atonal music was made according to their questionnaire responses in Phase 3, which had taken place 7 months after Phase 1 (Appendix 3): participants who had responded that their level of familiarity with atonal music had not been altered since Phase 1, were used as the baseline category. Participants who responded that their level of familiarity with atonal music had changed since Phase 1

and the result ranged from them being from 'Neither familiar nor unfamiliar' to 'Extremely unfamiliar' were placed in one category; participants who found their familiarity had changed to 'Familiar' or 'Extremely familiar' were placed in the second group. The hypothesis tested was that subjects who had increased their familiarity with atonal music between Phase 1 and Phase 3 would show more relative improvement both than subjects whose familiarity with atonal music had remained unchanged and than subjects who had responded that their level of familiarity had deteriorated. Multiple regression statistics are shown in Table 4.5:

$$\Delta_{Aba}$$

	B	SE B[a]	β
Constant	3.67	30.46	
Unchanged vs. familiar	293.63	85.40	.52***
Unchanged vs. neutral/unfamiliar	50.85	48.03	.12*

Note: R^2 = .27, p < .05.

* p > .05, *ns*

*** p ≤ .01, *s*

a.   Bootstrapped values. Significance for β is also based on bootstrap results.

Table 4.5: Multiple regression statistics for Δ_{Aba} , using familiarity with atonal music as the predictor variable.

The significance of the beta value for the contrast between the unchanged and the augmented familiarity group indicated that participants who had selected either the 'Familiar' or the 'Extremely familiar' answer in the questionnaire (Appendix 3) produced a statistically significant difference in their relative improvement, compared to all other participants. This finding was not consistent with predictors for the actual scores in the atonal performances in Phase 1: multiple regression using either the 1st or the 2nd performance of the Atonal piece, in either Phase 1 or Phase 3, as an outcome variable and familiarity with atonal music as the predictor, yielded non-significant results in all cases. This finding suggested that familiarity with atonal music was not, in itself, a significant predictor for performance in the Atonal piece; when, however,

familiarity increased to reach a level that the subject felt s/he was familiar with atonal music, performance showed more relative improvement between the 1st and the 2nd time. Many possible interpretations could explain this effect: one could be that, when the subject was not very familiar with atonal music, the 1st performance was the most accurate because it relied on short-term memory structures; when, however, familiarity increased, the subject was able to construct a more stable memory trace, leading to an improvement in the 2nd performance. This way, the change in familiarity level also signified a change in the mental procedure the subject had used to approach and break down the piece.

The non-significance of the contrast between the unchanged and the neutral/ unfamiliar group suggests that, although increased familiarity with atonal music seemed to provide a relative advantage in the memorised performance of the Atonal piece, non-familiarity did not necessarily entail a disadvantage in the ability to memorise atonal music.

This significant finding was only traced in the relative improvement between the 1st and 2nd performance: comparisons involving the 3rd performance, with familiarity with atonal music as the sole predictors, did not yield significant results, although for Δ_{Acb} the value of p for the overall model was just over the threshold of significance, at .051. The non-significance of Δ_Δ scores expressing relative improvement regarding the 3rd performance may, again, be interpreted as signposting the limits of short-term and long-term memory: familiarity with atonal music may have contributed towards the better retention of the piece after 15 minutes, but the lack of practice between the 2nd and 3rd performance – assuming that participants followed instructions – did not allow this effect to be upheld 24 hours later.

4.3.4. Δ_Δ scores: Frequencies

In addition to regression analyses, it is interesting to examine the frequency with which subjects showed relative improvement between performances in Phases 1 and 3. The frequencies displayed in Table 4.7 represent the percentage of participants whose Δ_Δ scores had a positive value, indicating that they showed relative improvement between the Phases for these performances; participants are categorised by Solfège use (I), which

is the distinction between the experimental (solfège) and the control (non-solfège) group and by Solfège use (E), in which solfège participants are only the members of the experimental group who used solfège in their 10-minute practice session in an audible way. Table 4.6 provides the absolute frequencies for the members in each group who successfully completed Phase 1 and Phase 3; AP participants are excluded from the count:

Table 4.6: Absolute frequencies

N	Tonal piece			Atonal piece		
	Δ_{Tba}	Δ_{Tca}	Δ_{Tcb}	Δ_{Aba}	Δ_{Aca}	Δ_{Acb}
Solfège (I)	12	12	12	12	12	12
Non-solfège (I)	15	14	14	15	12	12
Solfège (E)	8	8	8	7	7	7
Non-solfège (E)	19	18	18	20	17	17
Total	27	26	26	27	24	24

Table 4.6: Count of participants in each group. 2 AP participants are excluded from the count and one participant failed to complete Phase 3 of the experiment. One participant in Phase 3 failed to attend the 3rd performance of the Tonal piece and therefore no Δ_{Tca} and Δ_{Tcb} scores could be calculated; the same happened with 3 participants in the Atonal piece.

Table 4.7: Percentage frequencies of positive Δ_Δ scores

%	Tonal piece			Atonal piece		
	Δ_{Tba}	Δ_{Tca}	Δ_{Tcb}	Δ_{Aba}	Δ_{Aca}	Δ_{Acb}
Solfège (I)	41.7%	58.3%	50%	66.7%	58.3%	58.3%
Non-solfège (I)	26.7%	14.3%	57.1%	53.3%	75%	58.3%
Solfège (E)	12.5%	50%	62.5%	57.1%	71.4%	71.4%
Non-solfège (E)	42.1%	27.8%	50%	60%	64.7%	52.9%
Total	33.3%	34.6%	53.8%	59.3%	66.7%	58.3%

Table 4.7: Percentage of participants who showed relative improvement in performance between Phase 1 and Phase 3, yielding positive Δ_Δ values.

Percentage frequencies of positive relative change in scores between Phases 1 and 3 are very different for the experimental and the control group, regarding the Tonal piece: as shown in Table 4.7, only 26.7% of the control group improved their 2nd performance in Phase 3, as opposed to 41.7% of the experimental group. The difference becomes even larger in the relative improvement in the 3rd performance, where a mere 14.3% (2 subjects) of the control group performed better in Phase 3, as opposed to 58.3% (7 subjects) in the experimental group. The difference between the experimental and the control groups diminishes in the comparison of Phases between the 2nd and the 3rd performances, where approximately half of each group performed better in Phase 3; the explanation for this apparently strange effect is that the participants' 1st performance for the Tonal piece was considerably different than the next two, but did not change as much between the 2nd and the 3rd time. This finding is consistent with multiple regression analysis results, which revealed Solfège use (I) to be a significant factor in predicting relative improvement in performance between the 1st performance and either of the remaining two. Also consistent with non-significant regression results, percentage frequencies for positive Δ_Δ scores in the not-audible solfège (non-solfège *E*) and audible solfège use (solfège *E*) groups are discrepant between different performances: only

12.5%, or 1 out of 8 participants who audibly used solfège performed better in the 2nd performance in Phase 3, as opposed to 42.1% of participants who did not use solfège. This difference, however, was reversed in the 3rd performance, where 50% of solfège participants performed better than they had in the 1st, as opposed to 27.8% in the non-solfège group.

Possible explanations for this reverse effect should also contain an explanation for the discrepancy between the results of implicit and explicit solfège users; it is useful, at this point, to recapitulate what implicit solfège use entailed. Participants who belonged in the experimental group had completed an intensive solfège course and had successfully passed a solfège exam at the end of the course, which took place between Phase 1 and Phase 3 (see Chapter 2.2). It is important to note that no memorisation tasks, either with or without the use of solfège, were undertaken during the solfège lessons or for the solfège exam. All subjects in the experimental group were asked to use solfège in their memorisation, although the extent and the nature of the use was up to their discretion: they could either use it for certain passages, for the whole piece, singing, vocalising, or even mouthing the solfège syllables. The participants who had an audible solfège production lasting more than 1 minute in the 10-minute practice session, were placed in the explicit solfège use group; participants who could be heard using solfège for less than 1 minute were pooled with non-solfège participants, in the non-solfège (E) group. Out of 12 participants in the experimental group, 8 were placed in the explicit solfège use group; it is interesting to look at the profiles of the 4 individuals who did not make it in the group.

The 4 subjects who did not have audible solfège production of more than 1 minute in the recording of the practice session included 1 pianist, 2 string players and 1 wind instrument player. In the relevant question in the questionnaire, all but one of the string players reported they found that solfège had had a positive effect on their memorisation and further explained the nature of this effect; one string player found no effect. This suggested that, although not audible, they had, in fact, used solfège towards memorisation. The fact that they had not practiced solfège audibly can receive numerous explanations: apart from the wind player, who had the additional excuse of the explicit use of solfège automatically preventing simultaneous practice on the instrument, the most economical explanation would be that non-singers were more self-

conscious about singing solfège alone, especially since they knew they were being recorded by their solfège teacher. Although all participants were continuously asked to sing during solfège lessons, the singing was mostly done in groups or in pairs; only in the exam setting were the participants asked to sing by themselves. The fact that all singers in the experimental group used solfège audibly supports this explanation. Although it could be argued that these four participants had completed the questionnaire in the way they thought it was more 'appropriate', given the task and their solfège student–teacher relationship with the researcher, their competence in solfège, proved by their high marks in the solfège exam and the significant regression results of the Solfège use (I) predictor indicate that these participants had, in fact, used solfège and that this affected their memorisation performance.

Consistent with multiple regression analysis results, which yielded non-significant values for solfège use as predictor for relative improvement in the Atonal piece performances, positive Δ_Δ scores for the Atonal piece do not reveal any prominent differences between the solfège and the non-solfège groups (Table 4.7). Percentage differences between the explicit solfège groups are below 10 units for relative improvement between the $1^{st} – 2^{nd}$ and the $1^{st} – 3^{rd}$ performance; it seems, however, that the 3rd performance for the explicit solfège group was considerably better in Phase 3, as 71.4% of this group produced a relatively better 3rd performance. The difference with the explicit non-solfège group, for which the equivalent percentage was at 52.9%, suggested that, while for the solfège group the scores improved considerably across Phases between the 2nd and the 3rd performance, for the non-solfège group the scores remained relatively stable.

In the distinction between the experimental and the control group for relative improvement in Atonal performances, percentage frequencies reveal that the solfège group improved more in the 2nd performance, whereas the non-solfège group showed more relative improvement between the 1st and the 3rd performance in Phase 3. The percentage differences in relative improvement between groups in this comparison set and the explicit solfège use comparison groups are contradictory; as regression analysis did not provide any evidence regarding predictors, it is hard to draw any conclusions regarding solfège use and relative improvement in performance scores from this dataset. A final attempt to illuminate these results will be made in the subsequent section, where evidence from the questionnaire responses will be also examined.

Finally, it is interesting to see how Δ_Δ scores changed according to instrument, separated by solfège use (I). Regression analysis showed that instrument type was a significant predictor for relative improvement in the 3rd performance of the Tonal piece. This finding, in conjunction with findings from Experiment 1, led to the investigation of instrument type as a predictor for performance, both in Phase 1 and in Phase 3: multiple regression analyses showed that instrument type was a significant predictor (p < .05) for the 1st and 2nd performances of both the Tonal and the Atonal piece. An obvious problem with revealing any trends in the data distinguishing different instrumental types and solfège use is that the number of subjects in each group is extremely small, as shown in Table 4.8; in addition to the already small groups, some participants from the control group failed to attend the last experimental session, diminishing the numbers even further:

Number of participants categorised by instrument and Solfège use (I)

N	Tonal piece		Atonal piece	
	1st performance	3rd performance	1st performance	3rd performance
Singers – Solfège (I)	3	3	3	3
Singers – Non-solfège (I)	6	6	6	5
Monophonic – Solfège (I)	6	6	6	6
Monophonic – Non-solfège (I)	4	3	4	4
Pianists – Solfège (I)	3	3	3	3
Pianists – Non-solfège (I)	5	5	5	3

Table 4.8: Number of participants in each instrumental group in the experimental and control conditions. The same number of participants in each group completed the experiment in Phase 1 and 3.

Table 4.9 shows the percentage frequencies of participants in each group who had a positive Δ_Δ score, indicating a relative improvement between performances; because of the exceptionally small absolute frequencies in each group, positive Δ_Δ percentage frequencies are provided only for the difference between the 1st – 2nd and the 1st – 3rd performances of the Tonal and Atonal pieces, as these scores represent the more definitive differences between short-term and long-term retention:

Table 4.9: Percentage frequencies of positive Δ_Δ scores

categorised by instrumental group

%	Tonal piece		Atonal piece	
	Δ_{Tba}	Δ_{Tca}	Δ_{Aba}	Δ_{Aca}
Singers – Solfège (I)	0%	66.7%	66.7%	66.7%
Singers – Non-solfège (I)	0%	16.7%	66.7%	80%
Monophonic – Solfège (I)	50%	33.3%	66.7%	50%
Monophonic – Non-solfège (I)	25%	0%	25%	50%
Pianists – Solfège (I)	66.7%	66.7%	66.7%	66.7%
Pianists – Non-solfège (I)	40%	20%	60%	100%

Table 4.9: Percentage frequencies of participants in each instrumental group who yielded positive Δ_Δ values, indicating relative improvement in performance.

Table 4.9 shows that, for the Tonal piece, 2 out of the 3 singers who had used solfège performed better in the next-day performance of Phase 3, as opposed to just 1 out of the 6 singers who had not received any solfège instruction and had not used solfège. For the Atonal piece, on the other hand, there was no such trend, as the 3rd performance in Phase 3 improved for 2 out of 3 participants in the solfège group and 4 out of 5 participants in the control group.

Monophonic instrument players who belonged in the experimental group did not show a trend for relative improvement in either performance of the Tonal piece, whereas they showed more relative improvement than controls for the 2nd performance of the Atonal piece: 4 out of 6 solfège participants performed better in the 2nd performance of the Atonal piece in Phase 3 than they had done in Phase 1, whereas only 1 out of 4 control subjects improved in the 2nd performance of Phase 3.

Finally, pianists in the experimental group showed the same trend in relative improvement with singers: 2 out of 3 pianists who used solfège improved their 3rd performance of the Tonal piece in Phase 3, whereas the same improvement was displayed by only 1 out of the 5 pianists in the control group.

4.4. Discussion

4.4.1. Research question no. 1: Can adults learn and use solfège effectively?

The first hypothesis tested in the experiment stated that:

1. Musicians who learn solfège as adults are able to use it effectively in order to memorise music.

This statement was supported by statistical analyses, questionnaire responses and empirical findings. Transcriptions of the 10-minute practice sessions showed that most participants in the experimental group used solfège in an audible way during practice: 8 out of 12 participants used it audibly whilst practising the Tonal piece and 7 out 12 used it audibly in the Atonal piece. Multiple regression analysis results, however, suggested that extrovert, or explicit, practice was not necessary to affect performance; the significance of implicit solfège use, as well as questionnaire responses of the participants who had not sung solfège during the practice session, suggested that they had been able to use solfège note names to construct an alternative mental image of the piece. The responses of participants who, despite not having used solfège extrovertly, reported a positive effect of solfège use on their memorisation in the questionnaire, suggest that a version of the mechanism of *solidifying the musical trace via verbal encoding*, described in Chapter 1.2, had very possibly taken place: according to these subjects, the use of solfège in the Tonal piece enabled them to 'remember melodic shapes better', 'remember the note names in the difficult bits' and 'think about and remember intervals', while for the Atonal piece it helped them 'make sense of the

melody', 'concentrate on intervals than finger patterns', 'remember the beginning of each phrase', or simply 'remember note names along with [rhythmic] values'.

The feedback from the solfège lessons taking place in Phase 2 of the experiment, presented in section 2.2.4, also supported the hypothesis that solfège could be used effectively to learn new music; 78.95% of participants who felt they had learned solfège adequately stated they would use it in order to learn unfamiliar music (section 2.2.4). Participant responses in the questionnaires completed during Phase 3 also supported the view that solfège could be used effectively in music memorisation; participant responses in the relevant question are shown in Figure 4.1:
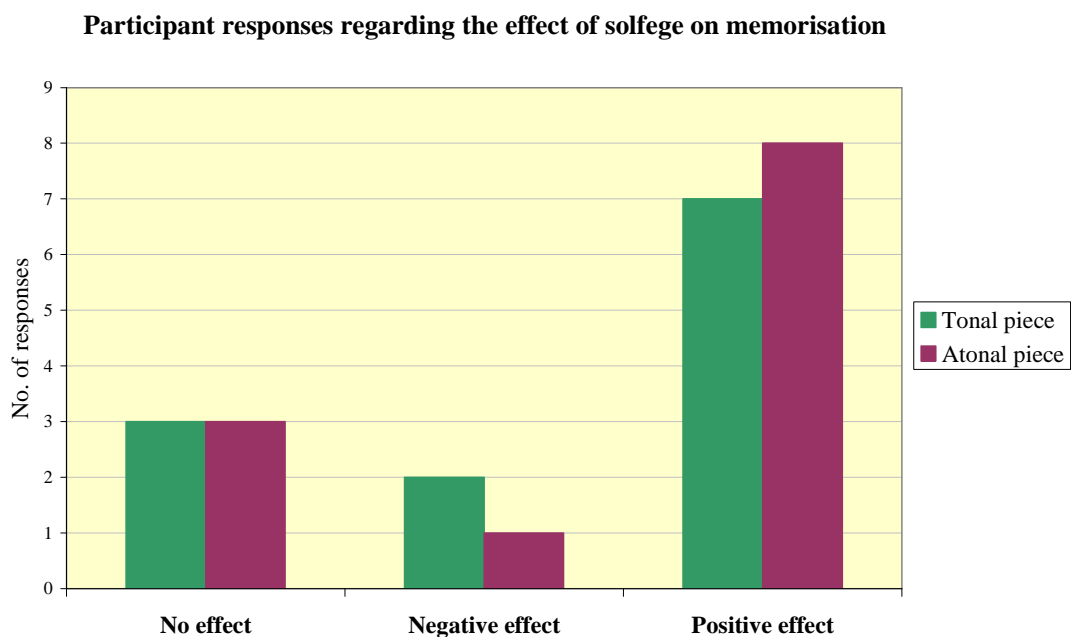
Participant responses regarding the effect of solfege on memorisation

Figure 4.1: Participant responses regarding the effect of solfège on their memorisation of the Tonal and the Atonal piece.

The bar chart in Figure 4.1 shows that most participants found a positive effect of solfège on their memorisation. The 2 participants reporting a negative effect were a flute player and a soprano; the flute player commented that she had found it confusing, because she had tried to memorise the piece singing solfège first, without using an instrument as a point of reference: as a result she had learnt incorrect pitches in singing,

so, when she played it, 'the patterns weren't right'. This was obviously a case of misuse of solfège, very possibly linked to the fact that the particular subject was performing on a wind instrument, meaning that it was more difficult to use it as an aid while learning to sing the piece. Although the participants performed the tasks in a practice room with a piano and were encouraged to use the piano if they needed to, none did; those who were less accurate in terms of pitch and intonation used their own instruments to learn the pitches. The fact that the particular participant had used neither a piano nor her flute as a type of 'tuner', rather than the use of solfège itself, was more likely the reason why she had found a negative effect on her performance. The responses of the remaining two participants reporting a negative effect of solfège on memorisation were less illuminating: a singer in the Tonal piece reported that solfège had 'thrown her concentration slightly', while a pianist in the Atonal piece found that solfège 'helps with tonal music more'. One of the most interesting explanations of the positive effect of solfège on the memorisation of the Atonal piece was provided by a soprano, who commented that solfège 'gives the piece an order when finding a melodic line is difficult'.

4.4.2. Research question no. 2: Does the use of solfège improve memorised performance?

The second hypothesis tested in the experiment stated that:

2. Learning and using solfège will improve performance in a music memorisation task compared to not using solfège in the same task.

Due to the violation of assumptions for parametric tests, this hypothesis was tested indirectly, using Δ_Δ scores in robust multiple regression analysis. Results supported the hypothesis for the Tonal piece, where implicit solfège use was found to be a significant predictor of relative improvement in the 3rd performance, in Phase 3. This finding signified that the last memorised performance of the Tonal piece, taking place 24 hours after the 1st, improved more, compared to the 1st performance, for subjects who had learnt and used solfège than for subjects in the control group. The relative improvement in the 3rd performance also provides support for the theory about solfège creating a more

stable memory trace, which is more easily stored and retrieved in long-term stores; the results of the present analysis justify further investigation of the exact mechanisms behind solfège encoding.

Although the hypothesis was supported for the Tonal piece, it had to be rejected for the Atonal: regression results showed that neither implicit nor explicit solfège use were significant predictors of relative improvement in performance scores. Percentage frequencies of positive Δ_Δ scores (Table 4.7) revealed that subjects who had audibly used solfège during practice had improved in their 3rd performance in Phase 3 by 18.5% more than participants who had not used solfège; as this finding, however, was not supported by regression statistics, further examination, using a larger sample, is required in order to reach a more definitive conclusion regarding the effect of solfège use on the memorisation of atonal music. It is worth noting, however, that questionnaire responses regarding the effect of solfège use on the Atonal piece, as depicted in Figure 4.1, did not differ much from responses regarding the Tonal piece, in which solfège use had been established as a significant predictor of performance. This suggests that, although analysis did not support the hypothesis for atonal music, the use of solfège in the Atonal piece provided participants with a 'psychological' advantage, since the use of solfège had, in most cases, made them feel more secure about the quality and effectiveness of their memorisation strategy.

4.4.3. Research question no. 3: Which instrumental group is benefited more by solfège use in memorisation?

The third hypothesis tested in the experiment stated that:

3. The positive effect of solfège utilisation in the memorisation process will be highest amongst singers and string players and lowest in piano players.

Due to the use of multiple regression analysis instead of ANOVA or MANOVA, post hoc tests and interaction graphs were not available in order to test this hypothesis; it was, therefore, tested indirectly, by cross-examining Δ_Δ frequencies and multiple regression statistics.

Multiple regression results provided partial support for the hypothesis, showing that singers were more likely to present relative improvement in their performance of the Tonal piece in Phase 3 than monophonic instrument players and pianists. Piano players, on the other hand, who were expected to perform worse than both singers and monophonic instrument players, appeared to deteriorate less than string and wind players compared to the singer group. The obvious problem with the interpretation of these statistics is that, although instrument type and solfège use were both significant predictors of relative improvement in the Tonal piece, this does not reveal the extent of interaction between the two parameters: it was not only singers from the experimental group who performed better in Phase 3, but all singers, from both groups. Restricted sample sizes in each instrumental group divided by solfège use (Table 4.9) render interpretations extremely tentative; it is worth noting, however, that 2 out of the 3 singers in the control group performed better in the next-day performance of the Tonal piece in Phase 3, as opposed to just 1 out of the 6 singers of the experimental group. The same trend was observed in the pianists' performances in the Tonal piece, whereas monophonic instrument players in the experimental group did not perform very differently than controls in terms of relative improvement. These frequencies suggest that, for the Tonal piece, an interaction between instrument type and solfège use may exist; further testing is required, however, in order to clarify the precise nature of the cross-effects.

Such a trend was not found in positive frequencies for Δ_Δ scores of the Atonal piece; only monophonic instrument players in the experimental group appeared to perform better in the 2nd performance in Phase 3 than controls. This lack of a prominent trend in Atonal piece data is consistent with multiple regression analysis results, in which neither instrument type nor solfège use were found to be significant predictors for performance.

4.4.4. Research question no. 4: Is solfège use more effective when used in atonal music?

The fourth hypothesis tested in the experiment stated that:

4. The positive effect of solfège utilisation will be stronger for atonal music than for tonal music.

This hypothesis was not supported by findings; on the contrary, solfège use appeared to be a significant predictor for relative improvement in the Tonal piece, whereas no such effect was found for the Atonal piece. This happened despite the fact that most solfège participants responded that they had found a positive effect of solfège use in their memorisation for the Atonal piece (Figure 4.1).

An explanation for this lack of a significant effect, despite the participants' intuitions, can be found in the recordings of the practice sessions for the Atonal piece. In many cases, the participants attempted to sight-sing the Atonal piece directly, without the help of an external point of pitch reference. This led to misconceptions regarding the musical material of the piece, which, in turn, obstructed the performance of the piece on the instrument. This, of course, could not have been the case with singers; the extremely small number of singers in the experimental group, however, does not allow comparisons which could verify or disprove this theory.

4.5. Summary

Despite the fact that the particularities of the dataset and the size of the sample did not allow definitive conclusions to be drawn for some of the experimental hypotheses, certain very interesting results arose. Solfège use appeared to improve subjects' performance in the Tonal piece, especially in the last performance, supporting the theory that solfège provides a stable encoding mechanism for musical material. Moreover, although this was not supported by statistical findings, participants found a

positive effect of solfège in their memorisation overall, both for the Tonal and for the Atonal piece; a possible explanation for the discrepancy between participants' intuitions and the results of analysis can be found in the way subjects approached atonal music singing. Nevertheless, solfège lessons elicited an enthusiastic response from participants, who, without any exceptions, stated in the solfège lessons questionnaires (Appendix 1) that solfège had improved some aspect of their musicianship. Analysis results provided concrete proof of the benefits of solfège use in music memorisation, especially for tonal music; moreover, the findings suggested that solfège use had an increased positive effect on performances relying on long-term memory. The next chapter will attempt to shed some light into the way solfège labels are stored in working memory, as a first step towards understanding the long-term effects of solfège use in memorisation.

Chapter 5

Experiment 3

The effect of pitch labels on music and verbal memory: A comparative study

5.0. Introduction

The use of Aospan and Arspan tests in the pilot study (Chapter 2.2) revealed that performance in these tests was not a reliable predictor of performance in the music memorisation tasks. This was believed to happen because operation and reading span tasks test a very broad set of skills, which factor in music memorisation performance in only a minor way (see section 2.2.3.2), compared to more domain-specific skills of musical expertise.

The benefit of using the specific operation and reading span tests was their established validity; devising new tasks would most probably carry the disadvantages of an ad hoc procedure, as the timeframe of the present research did not allow further piloting. On the other hand, even if the Aospan and Arspan tasks had yielded results which correlated well with the results of the music memorisation performances, the correlation would merely serve to support or reject the possibility of existence of an overriding relationship between working memory span and the specialized function of music memorisation, rather than shed light on possible hierarchical relationships between different components of working memory, or on musical and verbal working memory.

Since Aospan and Arspan proved inappropriate for the task at hand and the creation of music-specific working memory tasks became a necessity, we tried to devise tasks which would help define the function of verbal pitch labels within working memory. In accordance with the research aims (section 2.0.1.), the focus of the tests shifted from determining span to bringing out interactions between pitch tasks using pitch labels and verbal tasks. The following research question was formed:

If solfège names constitute verbal pitch labels with a unique musical dimension, will their memory trace bear more similarities to verbal elements (e.g. words), to musical units (e.g. pitches), or to a mixture of the two?

An attempt to answer this question should include the development of specialised tests for solfège names, which will be able to detect the existence of effects present in verbal or in non-verbal items, based on the lexicality impact on recall (Hulme et al, 1991;

Bourassa and Besner, 1994; Gathercole et al, 2001). Such a conclusive examination of solfège labels' attributes is beyond the scope and timeframe of the present research; certain preliminary tests, however, were developed in order to gain some insight regarding the nature of solfège labels.

The working memory tests described in the following chapters are based on music *reading*, rather than listening, and are therefore targeted to musicians only. The rationale behind this design was the following: the primary functional environment of solfège is, arguably, music reading; we would therefore expect all prominent characteristics of solfège to be manifested when reading, more so than when performing or listening to music. If solfège syllables display certain characteristics when tested within the music reading environment, then it will be plausible to research further into the same effects which can be possibly transferred to other functions; if, on the other hand, solfège syllables appear to lack certain characteristics when used in music reading, then it will be highly unlikely that they will display these characteristics in other circumstances. In addition, there is research suggesting that music reading can be regarded as an aspect of music perception (Sloboda, 1984; Wöllner et al, 2003).

In order to have an outside point of reference in the estimation of the effect of solfège use, solfège syllables were compared to letter names in the following tests. Letter names were chosen on the grounds of their popularity and their similarity with solfège syllables: if letter names and solfège syllables are interchangeable, their memory traces should have similar attributes.

The verbal and musical qualities of solfège memory traces, as described in the related research question, were investigated within two population samples: solfège and non-solfège. Solfège subjects used solfège syllables in the music recall task, while non-solfège subjects used letter names; other than the mode of recall and the language the test was presented in, the tests were identical. For all tests described in the following sections, the Aospan test (Unsworth et al, 2005) functioned as a model for the macro-organisation of the experiment. This decision was justified by two main factors:

1. Since one of the principal aims of the present study is to determine whether or not the mode of recall (syllables or letters) has an effect on memorisation, it is

       plausible to start testing this hypothesis by looking at whether the mode of recall affects span.

2. The Aospan test has been shown to be a reliable indicator of WM capacity, with the additional advantage of being fully automated, thus easily administered to participants. As the following tests do not aspire to provide anything more than tentative and complimentary evidence towards the main research question or, at most, lay the preliminary groundwork for more comprehensive research on this field, time-efficiency, both in terms of designing and running the experiments, was a crucial factor in selecting the particular tasks.

For the above reasons, the Aospan test was adapted to fit the aims of the present research by retaining the core concept of the test intact and changing the stimuli both in the main and in the distraction tasks. It is obvious that more comprehensive research is required in order to illuminate the field of verbally encoded musical pitch, including music-specific WM tests as well as brain-imaging, which will help locate the memory trace in different recall modes and define its attributes. The excellent work of V. J. Williamson (Williamson et al, 2006; Williamson et al, 2010) is in this vein, although her focus is considerably wider, testing music and language interactions in the general population rather than in expert musicians. The results of the following experiments provide a first step towards examining the way particular verbal labels affect music memorisation performance.

Another major restriction resulting from the selection of the particular type of behavioural tests used to investigate cognitive phenomena is that, naturally, the experimental manipulation can only control the type of output produced, rather than the process that leads to it. A prominent example is people who use their 'inner ear' abilities when asked to memorise music by simply reading it: these people will primarily use the auditory dimension of the stimulus to perceive, process and store it and will then 'translate' it into other types of output as required upon retrieval. In order to balance such effects, all tests were introduced to participants with extensive practice sections, which served the double purpose of familiarising them with the procedure and with the output required of them. Economy of resources suggests that subjects will use the same mode for storage as for recall, rather than 'translating' data between two

modes; this, of course, is only true provided that they will have the specified mode readily available in their cognitive mechanisms relating to music reading.

A total of 86 participants were tested in the following experiments: 45 were music students or staff members in the University of Hull and 41 were music students or staff members in the Aristotle University of Thessaloniki. All subjects tested in Hull had English as their first language and all music students from Thessaloniki had Greek as their first language; the tests were administered in English or in Greek, respectively. All Greek participants had been taught solfège during their childhood and were also following the solfège courses taught as part of their degree programme. Out of the 45 English participants, 9 had followed an intensive course in solfège training over the previous academic year and had passed the relevant exam (see Chapter 2.1): they constituted the 'taught solfège' group. It is important to note that the last systematic contact, i.e. in the form of lessons, the subjects in this group had with solfège had taken place 9 months before the testing. The remaining 36 English participants had never been taught solfège systematically and were not familiar with solfège names.

There were three core experiments (A, B and C), with three variations for the second and third experiments (B1, B2, B3 and C1, C2, C3), each one lasting approximately 10 minutes. Each participant was tested on any combination of three or four of the experiments in a single, 45-minute session; no participants completed all seven experiments: for example, participant no. 1 could complete experiments A, C1, B3, participant no. 2 experiments B1, B2, C3, and so on. This pseudo-random design was used in order to maximise the number of participants completing each test; as some of the tests featured identical stimuli in the memorisation or in the distraction tasks, there were limitations as to which tests a single participant could receive. The benefit of having more participants completing each experiment, of course, entailed that there was little control over the amount of participants completing *specific* series of experiments, which posed certain restrictions in data analysis; since, however, the present study had primarily an exploratory character, testing the biggest possible sample in each one of the experiments was considered extremely important. The exact durations of stimulus presentations used in the main study were all based on results and feedback from the pilot versions of the tests. In both countries, tests were administered to participants in

the computer lab facilities of the respective Music Departments; all tests were created using the E-Prime version 1.2.

After the end of the last experiment in the session, participants completed a short questionnaire, providing demographic information and background information regarding their music studies; the questionnaires were provided in English and in Greek, according to the country where the testing took place. The English version of the questionnaire can be found in the Appendix 3. The following sections contain an account of each of the three experiments (A, B, C) and their variations (B1, B2, B3, C1, C2, C3) separately; results will also be presented separately for each variation, with a summary of results at the end of each experiment (A, B, C). The combined results of all experiments and their variations will be discussed at the end of the chapter, in section 5.4.

5.1. Experiment A

The aim of this experiment was to provide a common ground for comparisons between the solfège group and the non-solfège group before testing for interference between verbal and musical data on recall; for this reason, it was designed as a purely music-reading task, without any parallel verbal interference tasks. The participants were asked to memorise and recall a four-note sequence, while performing a pitch-comparison music reading task.

The only difference between the tests received by the non-solfège, the solfège and the taught solfège group was the mode of recall (see Figure 5.1): non-solfège subjects were provided with a list of **letter note names**, whereas solfège and taught solfège subjects were provided with a list of **solfège note names**. For the non-solfège and the solfège subjects, the mode of recall provided was also the mode of recall they were most familiar with; for the taught solfège participants, however, solfège names, although familiar, were not expected to be used as automatically as letter names, especially taking into consideration the 9-month gap between the end of solfège lessons and the present tests. Taught solfège participants were thus expected to fall in two categories:

1. Participants performing a mental 'translation' between letter and solfège names, in order to recall material. These participants would be students whose solfège skills had, for various reasons, deteriorated over the 9-month gap and were, therefore, extremely uncomfortable using solfège names directly. They were expected to memorise stimuli thinking in letter names and then 'translate' them to solfège names when required. Although the experiment did not pose any time restrictions on recall, this 'mediated' use of solfège names was expected to lead to more confusions and mistakes than the direct use of either the letter or the solfège note names.

2. Participants trying to use solfège note names directly. These participants were expected to be students who had found solfège lessons the most beneficial (section 2.1.2) and had continued to use it in their personal practice, or students who had achieved high marks in the solfège exam and generally felt very confident about using solfège. Although these subjects had generally honed their solfège skills to a satisfactory level during the previous year, they could not be

considered fluent in the same sense solfège participants from Greece were fluent in using solfège names. For this reason, and given the time restrictions in stimulus presentation, these subjects were expected to get confused and make more mistakes in pitch names, compared to other subjects.

There was no way, of course, to determine the exact mode participants were using to memorise and recall pitch sequences; the only attempt to control the expected transition between the nominal and the functional stimulus, which is the stimulus as presented and the stimulus as encoded (Eysenck, 1977), was through providing mode-specific recall options (see section 5.1.1.3). Both in the case of 'translation' during recall, however, as well as in the case of attempts to directly memorise stimuli using solfège, taught solfège participants were expected to have a relative disadvantage compared to the other two groups of participants, when the use of solfège names was required. On the other hand, in cases when the mode of processing was not dictated by the experiment, i.e. in the distraction task, taught solfège subjects were expected to perform similarly to the other groups. For this reason, the hypotheses tested in the experiment were the following:

1. **Note span scores will be similar for both the non-solfège and the solfège group and lower for the taught solfège group.**

2. **The distribution of correct answers in the distraction task will be similar across the three groups.**

Due to practical issues in the design and conducting of the experiment, the size of the taught solfège group (N = 9) was significantly smaller than both the size of the solfège group (N = 41) and that of the non-solfège group (N = 36). This discrepancy was exacerbated further by the fact that not all participants undertook all 7 experiments (see section 5.1.1.2 for a report of the participants who completed Experiment A). Due to this anomaly in the sample, statistical comparison of the scores of the taught solfège group and the other two groups for the testing of Hypotheses 1, 2 and 3 was not possible; analysis is going to include results from the non-solfège and the solfège groups only and special mention will be made to the mean scores of the taught solfège participants.

5.1.1. Method

5.1.1.1. Design

The study employed an independent design, with one independent variable: solfège use, with three levels: solfège, taught solfège and non-solfège. The dependent variables were:

1. Note span, which was the total amount of notes remembered.
2. Total task correct, which was the total amount of correct answers given in the distraction task.

During the course of the analysis, other independent variables affecting performance were examined, such as first-study instrument category.

5.1.1.2. Participants

55 musicians in total completed the task, of which 27 were English and 28 were Greek. 5 of the English participants had received solfège training over the previous academic year and were therefore familiar with solfège note names; these participants completed the English version of the solfège test. The remaining 22 English participants completed the task using letter names, whereas all Greek participants completed the task using solfège names.

The English group included 17 men and 10 women, with an average age of 22.96 years; 25 participants were undergraduate or postgraduate music students in the University of Hull and 2 were staff members in the department of Drama and Music. 23 of the English participants stated they did not have AP and 4 stated they did not know whether or not they had AP. There were 9 singers, 2 string players, 3 wind players, 8 pianists, 4 guitar players and 1 percussionist in the English group.

The Greek group included 9 men and 19 women, with an average age of 21.57 years; 27 participants were undergraduate or postgraduate music students in the Aristotle

University of Thessaloniki and one was a member of staff in the Department of Music Studies. 1 Greek participant possessed AP and another one stated she did not know whether or not she had AP. There were 4 singers, 2 string players, 2 wind players, 17 pianists and 2 guitar players in the Greek group.

5.1.1.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli consisted of four-crotchet sequences in treble clef, presented in a single staff for 5 seconds. The 4 notes did not contain repetitions and did not have an obvious, straight-forward tonal connection, although they were always natural notes, without accidentals; each stimulus was presented as one 4/4 bar. The presentation of the stimuli was visual only, with no accompanying sound. An example of a 4-note sequence to be memorised is shown in Music Example 5.1:

Music Example 5.1: One of the musical stimuli participants were asked to memorise.

For the recall of the sequence, a screen appeared, which contained a list of all notes with their letter or solfège names, according to whether the subject belonged in the solfège, the taught solfège, or the non-solfège group; a green box was placed next to each note. The subject was instructed to choose the correct notes in the correct order by clicking on the relevant boxes; selected boxes turned red and acquired a number between 1–4, to indicate the order in which they had been clicked on. In addition to the seven note names, a set of four 'Blank' options was available: in case a participant forgot or was dubious about the pitch of a note, they could click next to one of the 'Blank' options, in the respective order they thought that note occurred, and then proceed with the rest of the sequence. A screenshot of a recall screen is shown in Figure 5.1:
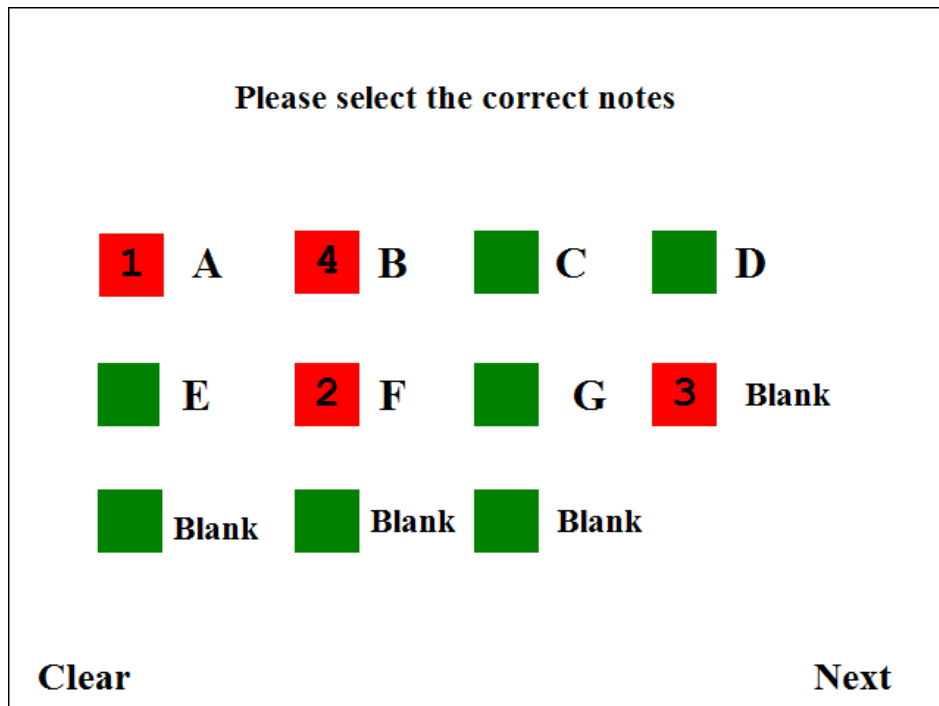
Figure 5.1: A recall screen for the memorisation task using letter names. In this instance, the subject could not recall the 3rd note of the sequence, therefore clicked one of the 'Blank' options in its place.

A 'Clear' button was also provided, which could be used in case the subject made a mistake and needed to undo all previous selections. There was no time restriction in the recall screen; participants could spend as much time as they wished putting their response together. Once the subject finished their response, they clicked on 'Next' to proceed with the rest of the experiment. Before the commencement of the next task, a feedback screen appeared for 2 seconds, which informed the subject of the exact number of notes s/he had remembered correctly. An example of a feedback screen is shown in Figure 5.2:

You scored 2 out of 4
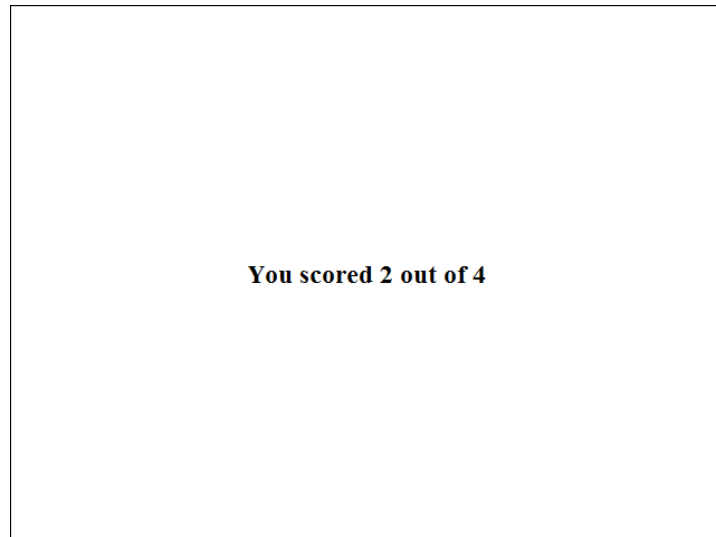
Figure 5.2: A feedback screen for the memorisation task.

5.1.1.4. Materials and Procedure: Pitch comparison task stimuli

Between the presentation of the memorisation stimulus and its recall, participants were presented with a series of note pairs. These pairs included one note in treble and one note in bass clef, in random order; in this task, too, the presentation of stimuli was visual only. Each stimulus stayed on screen for 4 seconds, during which participants had to decide which note of the two had the highest pitch. No accidentals were used; the reason both bass and treble clefs were used in this task was to prevent participants from turning the pitch task into a spatial positioning task, as, naturally, the higher note of the two in the same clef is also the higher note on the staff. An example of the high–low comparison task is shown in Music Example 5.2:

Music Example 5.2: One of the high–low comparison tasks.

The participants were instructed to select which note, in order of staff presentation, they thought was the highest in pitch by pressing 1 or 2 respectively in the numeric keypad. If the participant thought that the first note (middle C in the example) was higher, they were instructed to press 1 on the numeric keypad; if they thought the second note (F4 in the example) was higher, they had to press 2. A feedback screen followed each stimulus presentation, which lasted 1 second and contained one of the following three types of feedback, according to the subject's response: 'correct', 'incorrect' or 'no response', in the event the subject had failed to press any key during the 4 seconds of stimulus presentation. The high–low tasks were presented in sets of 2–7 items, featuring different pairs of notes each time. The importance of performing the task as quickly and as accurately as possible was stressed throughout the experiment.

5.1.1.5. General procedure

The whole of the experiment was mouse-and-keyboard driven, without any need for intervention from the experimenter. The experimenter, however, remained in the same room and was available to provide any help required during the practice trials; subjects were not allowed to talk or ask any questions during the main experiment.

Before the beginning of the main core of the experiment, the participant read the on-line instructions and had the chance to perform practice trials, in order to get familiarised with the procedure. As in the Aospan test, the practice trials were broken down in three parts: in the first part, the subject practised only the memorisation task; in the second part, the subject practised the high–low task separately and in the third part, the subject combined the memorisation and high–low tasks, in a simulation of the experimental procedure.

The core procedure comprised three experimental sets; in each set, the memorisation stimulus was presented first, the high–low task set followed and the recall screen for the memorisation task appeared at the end. The number of high-low tasks in each experimental set varied between 2 and 7 tasks, while the overall number of high-low tasks, across all three experimental set was between 11 and 15 tasks. A sample procedure of a single experimental set is shown in Figure 5.3:
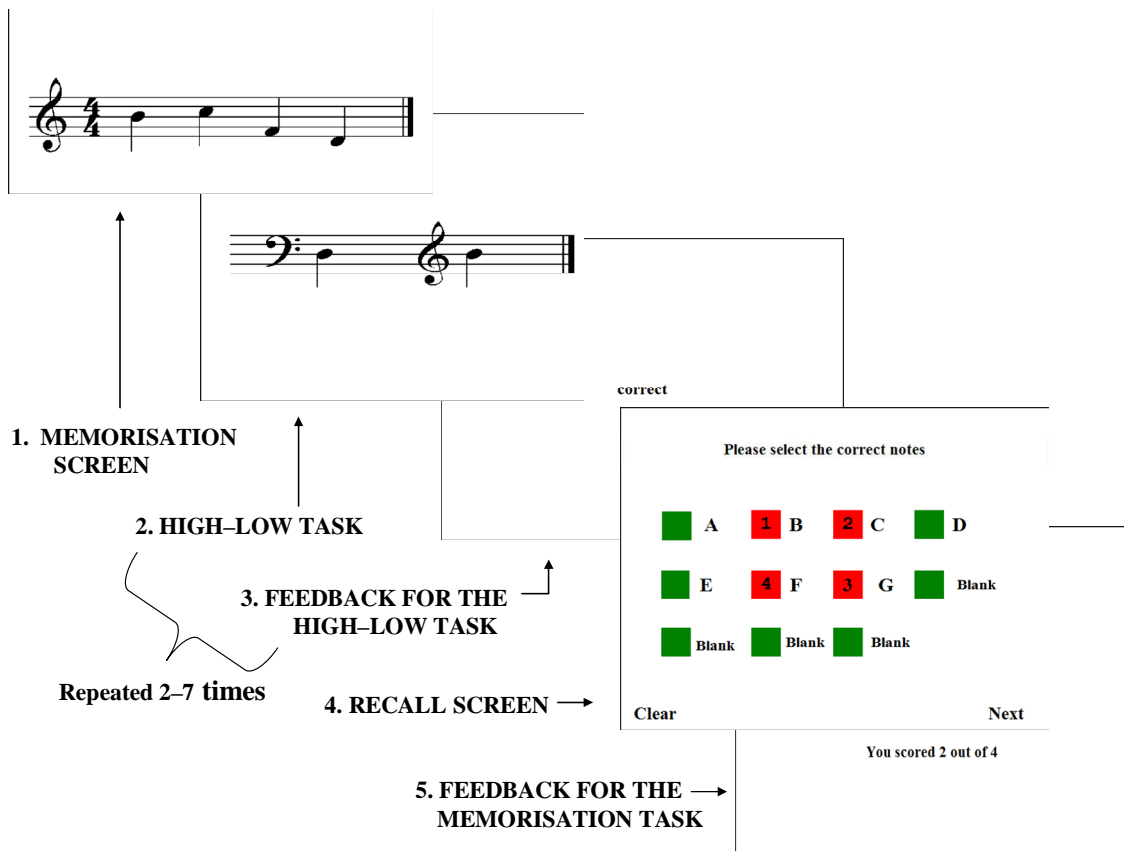
Figure 5.3: A schematic presentation of the procedure followed in each experimental set, based on the model in Unsworth et al, 2005. Note that stages 2 and 3 were repeated between 2 and 7 times. The core procedure is identical to the one of Aospan tests (Unsworth et al, 2005).

After the end of the third experimental set, the participant was informed that the experiment was over and was instructed to inform the researcher. Data logging from the experiment allowed calculation of the following scores:

- **Note span**, equal to the total number of notes remembered correctly in the correct position, across all 3 experimental sets. The maximum score was 12, which is the sum of all notes in the three 4-note sequences used in the memorisation tasks.

- **Total errors**, equal to the total number of errors made in the high–low distraction task. This score is further broken down in two types of errors:

  - **Speed errors**, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

- ■ **Accuracy errors**, which is the total number of errors resulting from the subject answering incorrectly to the high–low task.

In addition to these scores, the number of *correct* answers given in the distraction task was also used in the analysis; it was calculated by subtracting the Total errors score from the total number of high–low tasks performed throughout the 3 experimental sets. This was the **Total task correct** score and was calculated only as a percentage score.

All the above scores were also calculated in percentage values. Before the start of the experimental procedure, participants were reminded by the researcher that it was extremely important to perform not only the memorisation task accurately, but to complete the pitch comparison task as quickly and as accurately as possible. This additional comment was made in order to be absolutely clear to all participants that the pitch comparison task was not, in any way, secondary to the memorisation task; it was explained that it was as essential to produce correct answers in the pitch comparison task as it was in the memorisation task.

5.1.2. Results

Table 5.1 displays the descriptive statistics for the memorisation and the distraction task scores separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	**Median**	**SD**	**Min.**	**Max.**	**Skew**	**Kurtosis**
Note span	Solfège	11.32	12	1.39	7	12	- 2.23	4.08
	Taught solfège	9	10	4.12	2	12	- 1.71	3.09
	Non-solfège	10.36	12	2.54	4	12	- 1.43	0.89
Note span (%)	Solfège	94.35	100	11.58	58.33	100	- 2.23	4.08
	Taught solfège	75	83.33	34.36	16.67	100	- 1.71	3.09
	Non-solfège	86.36	100	21.14	33.33	100	- 1.43	0.89
Total high–low errors	Solfège	2.36	1	3.12	0	13	2.17	4.97
	Taught solfège	3.40	4	1.82	1	5	- 0.57	- 2.23
	Non-solfège	2.73	2.50	2.59	0	8	1.04	0.25
Speed errors	Solfège	0.50	0	2.46	0	13	5.24	27.63
	Taught solfège	0.40	0	0.89	0	2	2.24	5
	Non-solfège	0.32	0	0.78	0	3	2.66	6.86
Accuracy errors	Solfège	1.86	1	2.35	0	10	2	4.63
	Taught solfège	3	3	1.58	1	5	0	- 1.20
	Non-solfège	2.41	2	2.42	0	8	1.08	0.69
Total task correct (%)	Solfège	83.34	91.29	22.45	0	100	- 2.44	6.84
	Taught solfège	74.09	69.23	15.73	54.55	93.33	0.12	- 1.63
	Non-solfège	80.73	83.33	18.39	38.46	100	- 1.07	0.29

Table 5.1: Descriptive statistics for Experiment A scores, by solfège knowledge group. It is important to note the difference in group sizes: for the solfège and the non-solfège group N = 28 and N = 22, respectively, whereas for the taught solfège group N = 5.

A value which raised some concern regarding the quality of performance in the distraction task was the zero Total task correct score in the minimum value for the

solfège group, combined with the maximum number of speed errors in the same group: these numbers suggested that there was at least one solfège subject who had failed to complete the distraction task in parallel with the memorisation. Box plots were examined and revealed that both scores belonged to the same subject who, in an experimental procedure comprising 13 distraction tasks (4, 5 and 4 in each set, respectively), had failed to produce any response within the 4-second timeframe. As the purpose of the experiment was to test performance on the memorisation task monitoring the interference caused by the parallel task, data from this participant were excluded from all further analyses; for all subsequent analyses, the number of subjects in the solfège group is $N = 27$, unless indicated otherwise.

After the removal of the outlier, Note span and Total task correct scores were plotted, both collectively and separately for the solfège and non-solfège groups. It is important to remember that Note span could receive values between 0–12, reflecting the exact amount of notes recalled in the correct position across the three experimental sets, and that Total task correct is a score expressing the percentage of correct answers a subject got in the distraction task. In both cases, scores appeared to have significantly non-normal distributions, as shown in Table 5.2:

Tests of Normality

	Note span			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.419	.704	.614	.797	.862	.859
df	27	22	54	27	22	54
Sig.	.000	.000	.000	.000	.006	.000

Table 5.2: The Shapiro-Wilk test was preferred to the Kolmogorov-Smirnov because of the relatively small sample sizes. Test statistics are shown for the distribution of scores in the two groups separately as well as in the total sample, including solfège, taught solfège and non-solfège participants. All distributions were found significantly non-normal at $p < .01$.

Because of the violation of the normality assumption and the unequal sample sizes, non-parametric tests were preferred to a one-way ANOVA or MANOVA for data analysis:

Mann-Whitney as well as Kolmogorov-Smirnov Z tests were conducted, as the latter tends to have better power in analysing relative samples (Field, 2009). The Moses test of Extreme Reactions was also used to check differences in the range of scores between the two groups.

For the Note span scores, both the Mann-Whitney and the Kolmogorov-Smirnov test statistics were found non-significant at $p > .05$, supporting the null hypothesis that the distribution of Note span scores was the same across the solfège and non-solfège groups. The Moses test of Extreme Reactions, on the other hand, was found significant at $p < .01$, indicating that the range of scores was significantly different between the groups. Note span scores were examined in terms of score occurrence frequency and outliers and the significant Moses test was attributed to the fact that 2 participants in the non-solfège group (9.1% of the cases) had a score of 4 and 5, respectively, while the remaining non-solfège participants had scored between 7 and 12; in the solfège group, all participants had scored between 7–12. The significant Moses test was, thus, attributed to the relatively small size of the sample, in which the 2 outliers skewed the non-solfège results significantly.

For the Total task correct scores, the Mann-Whitney and the Kolmogorov-Smirnov test statistics were also found non-significant at $p > .05$, supporting the null hypothesis that the distribution of correct answers in the high–low task was the same across the solfège and non-solfège groups. The Moses test of Extreme Reactions was also found non-significant at $p > .05$, showing that the range of Total task correct scores was the same across groups.

Non-parametric tests supported the experimental hypotheses regarding similarity in the distributions of Note span and Total task correct scores. As the extremely limited size of the taught solfège sample ($N = 5$) rendered statistical analysis for this subgroup impossible, comparisons were made by looking at the outliers and the mean scores of the three groups. The exact Note span scores that the 5 subjects in the taught solfège group had achieved are shown in Figure 5.4:
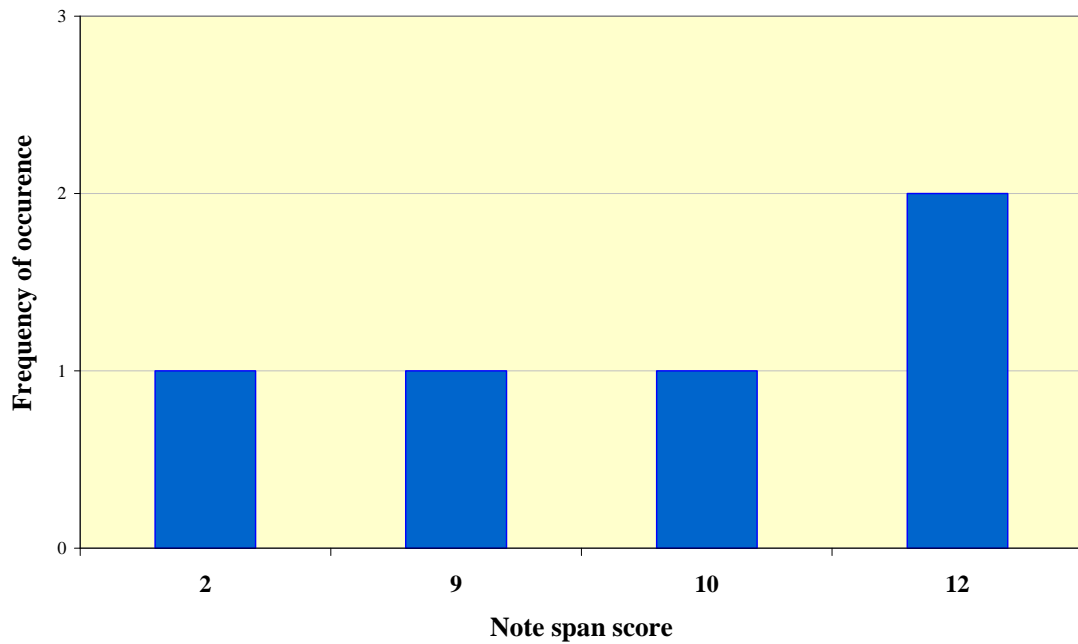
Figure 5.4: The distribution of Note span scores in the Taught solfège group. 1 subject recalled 2 notes in the correct position, 1 subject recalled 9 notes in the correct position, 1 subject recalled 10 notes in the correct position and 2 subjects recalled 12 notes in the correct position.

According to the first experimental hypothesis, taught solfège participants were expected to have more trouble naming the recalled pitches accurately, due to an assumed 'translation' process taking place. This, however, could not be confirmed by the data from the 5 subjects who completed Experiment A: although one subject yielded a note span of 2, which was the lowest span across the three groups, the remaining four taught solfège subjects had scored between 9–12, producing a mean score of 10.75, which was between the mean scores of 10.36 for the non-solfège and 11.32 for the solfège group. These figures did not reveal any tendency amongst taught solfège participants to perform worse in the music memorisation task.

According to the second experimental hypothesis, taught solfège participants were expected to perform in a similar manner to the other two groups in the distraction, high–low task. The mean scores for the taught solfège, solfège and non-solfège groups are provided in Table 5.3:

Total task correct (%) scores

	Mean	SE
Solfège	86.43	3.02
Non-solfège	80.73	3.92
Taught solfège	74.09	7.03

Table 5.3: Mean Total task correct scores for the three solfège groups.

The mean Total task correct score for the taught solfège group was quite lower compared to the respective mean of the other two groups and had a larger associated standard error; score frequencies were checked and revealed that this difference could not be attributed to outliers: three out of the five taught solfège subjects had a Total task correct score below 70, whereas the equivalent numbers for the other groups were 2 out of 27 (7.4%) for the solfège group and 4 out of 22 (18.2%) for the non-solfège group. Although this difference in mean scores could be a result of small sample sizes in general, it was enough to justify further scrutiny of the scores.

If mean Total task correct scores were considered to reveal a tendency amongst taught solfège participants to make more mistakes in the distraction task, one possible explanation for this could be that the proposed 'translation' process described in Chapter 5.1, rather than having a direct detrimental effect on the memorisation task, it affected the distraction task negatively through a trade-off of attention. This would mean that, because the particular five subjects did not rely on an automated process for encoding and keeping the trace of the 4-crotchet sequence active in working memory, they had to devote more cognitive resources in the memorisation/recall task rather than the distraction task, leading to poor performance in the latter. An obvious problem with this theory is that it implies that the 'translation' of letter names into solfège names took place much earlier on than the recall screen; for the attention trade-off to exist, the translation would have to take place during the 5 seconds of stimulus presentation. Although this may sound counter-intuitive in terms of time management, the fact that subjects had the chance to perform at least three practice trials before the actual experiment meant that subjects were anticipating that the mode of recall would be based on solfège-names; this knowledge could very possibly have led subjects, especially the

more confident ones, to memorise the sequence using solfège names. Informal feedback given spontaneously to the researcher at the end of the overall experimental session, regarding the process taught solfège participants had followed to memorise the sequences, was circumstantial and contradictory, so that it did not help to clarify the situation. The number of speed and accuracy errors was also checked, as more speed errors compared to accuracy errors would provide support for the theory, signifying that taught solfège participants required more time to perform the distraction task; speed errors, however, had only been made by one taught solfège subject and occurred, in general, extremely rarely in the sample, compared to accuracy errors.

Despite the sample size not allowing any definitive conclusions to be drawn regarding a potential negative effect of the taught solfège condition on the performance of the high–low task, the examination of scores revealed the need to monitor performance on the memorisation task *in relation to* performance on the distraction task. This meant that it had to be ensured that excellent performance on the memorisation task was not accompanied by poor performance on the distraction task, signifying that participants tended to ignore the latter in order to perform well on the former. Correlation of the two scores was examined both for the whole sample and for the solfège and non-solfège groups separately. For the sum of results, Note span was found significantly correlated to Total task correct, $\tau = .29$, $p < .01$. For the non-solfège group, the same correlation was found to exist, $\tau = .39$, $p < .05$; for the solfège group, however, Note span was not correlated to performance on the high–low task, $\tau = .17$, $p > .05$. These results suggested that participants from the non-solfège group who performed well on the memorisation task were more likely to have performed well on the high–low task, too, whereas this tendency did not appear in solfège participants. The fact that the tendency appeared in the sum of results, including solfège, taught solfège and non-solfège participant scores, was considered to be a transfer effect from the significance of the effect in the non-solfège group. The positive correlation suggested that good performance on the distraction task did not equate to poor memorisation performance but, on the contrary, subjects who performed well in one task tended to also perform well in the other task; this, in turn, indicated that the tendency of taught solfège participants to produce more errors than other groups in the distraction task, whilst keeping memorisation scores relatively intact, was an occurrence isolated in the particular group. This occurrence could be explained either by sample size, meaning that this was a chance finding, or by

the translation process, described previously, interfering with attention resources; testing a much larger sample would be the only plausible way to support either of the two theories.

Having established that solfège knowledge most probably did not affect Note span and Total task correct scores, it is interesting to examine other possible factors affecting performance. Data analysis from Experiment 1 (see Chapters 3.3 and 3.4) suggested that the factor that mostly affected memorisation performance was the instrument performed. Although the present experiment was different, in that the first-study instrument was not directly involved in the completion of the experiment, it was considered worthwhile examining whether or not there was a link between participants' first-study instrument and their Note span and Total task correct scores.

The total sample minus the outlier (N = 54) included 12 singers, 4 string, 5 wind, 26 piano, 6 guitar players and one drummer. As instrumental categories were quite many and very discrepant in size, further categorisation was required in order to conduct meaningful analysis. Experiment 1 results (Chapter 3.3) had shown that monophonic, i.e. string and wind, instrument players produced similar patterns of results, so they were pooled together in one group; the same was done for the purpose of the current analysis. Guitar players formed a separate category, although it is important to note that three of them were classical guitarists and three were jazz and pop guitar players. The drummer was also added to the guitar group, as more relevant to the jazz and pop guitar group than to any other instrumental group of those available. Although the categorisation of a drummer with guitarists may seem odd, it is worth noting that in the present experiment the precise nature of the instrument was not expected to affect as much as the general music-reading context in which the instrument is normally performed. The resulting instrumental categories were:

a. Singers (N = 12)
b. Monophonic instrument players (N = 9)
c. Pianists (N = 26)
d. Guitars and drummer (N = 7).

For Note span scores, the Kruskal-Wallis test statistic was found significant at $p < .05$, indicating that the first-study instrument category affected the distribution of scores. In

order to discover which group, or groups, in particular caused the change, robust multiple regression was preferred to post-hoc tests for the Kruskal-Wallis test, in order to avoid inflated Type I and Type II errors. Multiple regression analysis using bootstrap samples was conducted in the data, using Note span score as the outcome variable and the four instrumental categories as predictors; as in Experiment 1, singers were selected as the baseline category, against which the other three categories were compared. Multiple regression analysis results are shown in Table 5.4:

Note span scores

	B	**SE B**[a]	**β**
Constant	10.85	.47	
Singers vs. monophonic	.27	.78	.04*
Singers vs. pianists	.50	.56	.11*
Singers vs. guitar-drums	− 3.13	1.45	− .46**

$R^2 = .27$, p = .001

* p > .05 (2-tailed), *ns*

** p < .05 (2-tailed), *s*

a. Bootstrapped values. Significance for β is also based on bootstrap results.

Table 5.4: Multiple regression statistics for Note span scores.

The value of R^2 change indicated that the predictors affected the model significantly; beta values, however, indicated that the only significant predictor was the difference between the singer and the guitar-drums group: participants in this group appeared to have significantly lower Note span scores. It is worth noting that the associated standard error was much higher for this predictor group than it was for the difference between singers and monophonic instrument players or singers and pianists; this suggested that the guitar-drums group showed more internal heterogeneity than the other groups. This was attributed to the inherent heterogeneity of participants in the group; although the categorisation was considered optimal for the analysis given the sample restrictions, the group still contained a rather arbitrary selection of participants. Due to skepticism ensuing from this sampling flaw, score frequencies, mean scores and outliers in groups were examined. Mean scores for each instrumental group are displayed in Figure 5.5:
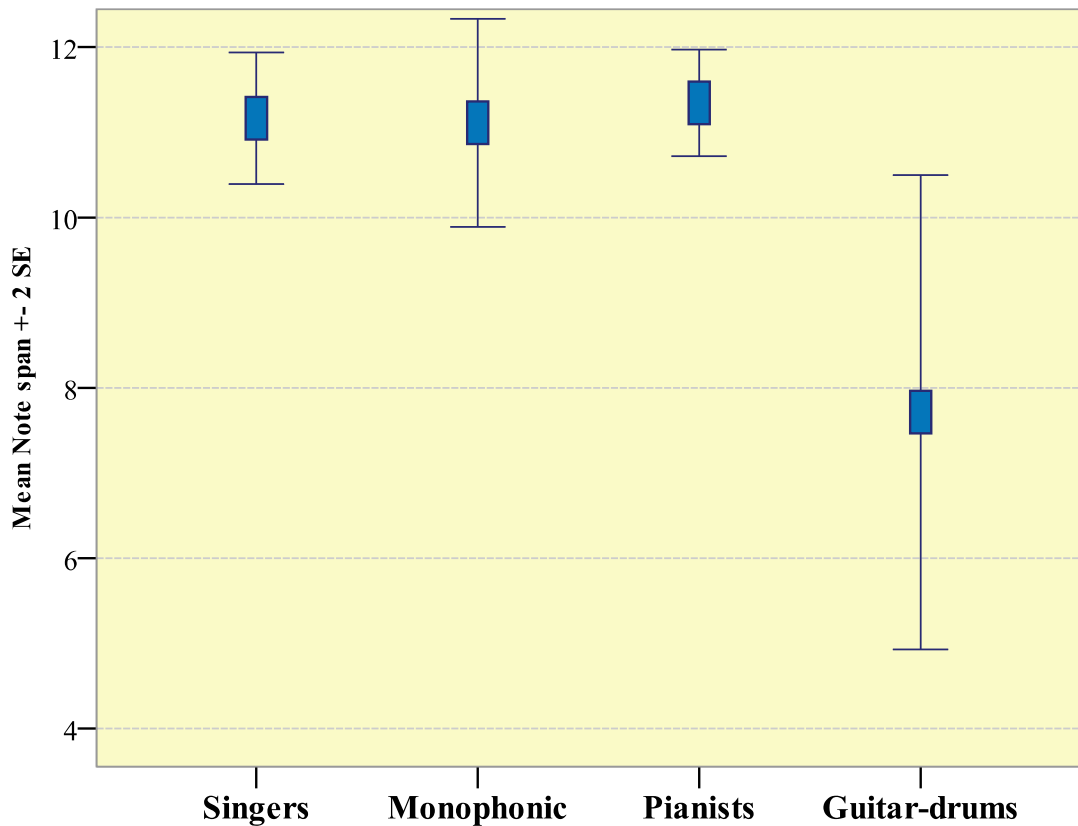
Figure 5.5: Mean scores for the four instrumental groups: for singers (N = 12) M = 10.85, for monophonic instrument players (string and wind players) (N = 9) M = 11.11, for pianists (N = 26) M = 11.35 and for the guitar-drums group (N = 7) M = 7.71. Error bars represent the mean ± 2 standard errors; standard errors were preferred over confidence intervals in this case because of the extremely limited size of the taught solfège sample.

The mean scores of the first three groups in Figure 5.5 were very close, whereas the fourth group displayed a significantly lower mean. Score frequencies were examined and revealed that 2 of the 7 participants in the last instrumental group had achieved a score of 2 and 4, respectively; the remaining scores were between 7–12. All scores in the singer and monophonic instrument groups were found between 7–12; in the pianist group, 1 subject had a Note span of 5, while the remaining 25 subjects had scored between 8-12. It was, therefore, obvious that the low mean score in the guitar-drums group and the subsequent significant regression results could be attributed to the two lowest scores in the group; further examination revealed that the lowest scores belonged

to a non-classical guitarist and a drummer. A possible explanation for the low scores could be that the lack of practice using sheet music caused these two subjects to perform poorly in the music memorisation task; it is worth noting that the guitar player was the only guitarist in the group who had never received formal instrumental instruction, either on guitar or on any other instrument: this could provide a plausible explanation for the difference between this subject's performance and the other guitarists in the group. The drummer, on the other hand, had received guitar lessons for 6 years, but it was not clarified whether the lessons were on classical, acoustic or electric guitar; in the very likely event that these guitar lessons did not include reading staff notation, the apparent difficulty displayed by the particular subject in remembering note sequences could be attributed to lack of familiarity with staff notation, as for the low-span guitar player. This theory supported the notion that the two subjects should be perceived as outliers; testing a much larger sample, however, would be necessary in order to provide further support for this view.

For the Total task correct scores, the Kruskal-Wallis test statistic was found non-significant at $p > .05$, suggesting that different instrumental groups (singers, strings, winds, pianists, guitarists and drummer) produced similar distributions of correct answers. This result was followed up with robust multiple regression, using the same predictors as for the Note span model: singers were used as a baseline group and their performance in the distraction task was compared to the performance of monophonic instrument players, pianists and pooled performances from the guitar-drums group. The effect of the predictors in the model was found non-significant at $p > .05$, providing further support for the fact that differences in the first-study instrument did not yield differences in performance of the high–low task.

As the high–low task included stimuli presented in treble and bass clefs, a meaningful comparison was considered to be the one between instrumentalists and singers who normally use bass clef and those who do not; the Mann-Whitney and the Kolmogorov-Smirnov Z test statistics appeared non-significant at $p > .05$, indicating that instrumentalists and singers who normally use bass clef, such as cellists, pianists, tuba players and baritones, were not more likely to perform better in the distraction task than those who normally use treble or alto clefs when performing music.

5.1.3. Summary

The first hypothesis stated that:

> Note span scores will be similar for both the non-solfège and the solfège group and lower for the taught solfège group.

Non-parametric test statistics supported the first part of the hypothesis: the distribution of Note span scores was found to be the same for the solfège and the non-solfège groups. The second half of the hypothesis, however, could not be supported by the data collected in the present experiment: the five participants in the taught solfège group did not appear to have big differences in their performance compared to the other two groups. Nevertheless, more extensive testing is required in order to reach a definitive conclusion regarding the distribution of scores in the taught solfège group.

The second hypothesis stated that:

> The distribution of correct answers in the distraction task will be similar across the three groups.

Once more, the hypothesis was only partly supported; although solfège and non-solfège participants yielded similar Total task correct scores, data from the taught solfège group, however sparse, raised some concerns regarding the homogeneity of errors across groups. In order to examine this effect, a theory was put forward explaining low performance scores on the distraction task as a result of a trade-off effect in attention between the memorisation and the distraction tasks. This effect was not found to be generalised across the three groups, therefore the theory attributing it to a special behaviour of the taught solfège group was not disproved; running the tests in a much larger scale, however, would be essential in providing any conclusive evidence supporting or rejecting this theory.

Apart from solfège knowledge, other predictors were examined, which could potentially affect Note span and Total task correct scores. The guitarists and drummer group appeared to yield significantly poorer performances for Note span; the co-examination of score frequencies and questionnaire data, however, revealed that this was more likely

to be an effect of lack of staff notation practice, rather than a consequence of the instrument performed itself. Nevertheless, further testing using a more comprehensive sample would be required in order to illuminate this point. Regarding Total task correct scores, results appeared unaffected both by instrumental category and by the lack of performance practice on bass clef. Combined with the fact that Total task correct scores were significantly correlated with Note span scores in the sample, results suggested that performance in the distraction task was mainly affected by individual differences in music reading and / or musical experience than by demographic factors.

5.2. Experiment B

The aim of this series of experiments was to compare the effect of a pitch-comparison task on visual verbal memory in solfège and non-solfège musicians. In order to achieve this, visual verbal memory was tested on three levels:

1. **Phonemic level:** In the verbal memorisation task, subjects were asked to memorise a series of 2–7 letters.

2. **Morphemic level:** In the verbal memorisation task, subjects were asked to memorise a series of 2–7 nonsense syllables.

3. **Lexical level:** In the verbal memorisation task, subjects were asked to memorise a series of 2–7 lexical morphemes, i.e. monosyllabic words.

In all three variants of the experiment, a verbal memorisation task was run in parallel with a pitch discrimination task.

5.2.1. Experiment B1: Musical interference in a phonemic memorisation task

In Experiment B1, participants were asked to memorise letter sequences while simultaneously performing a visual pitch discrimination task, identical to the pitch discrimination task in Experiment A (Chapter 5.1.1.4). The use of letter names to identify pitches was expected to interfere with the memorisation of letters and have a detrimental effect on the performance of the memorisation task. An important restriction in the design of the experiment was that there could be no specified mode for the performance of the distraction task; subjects could be separating pitches into higher and lower by thinking about their names, by trying to reproduce them mentally, imagining their spatial position on their instrument or voice, by using their 'inner ear' abilities, or by using any combination of the above methods. The resulting ambiguity entailed that, even if the hypotheses were supported by data analysis, further research would be required in order to illuminate the nature of the effect; if, on the other hand, the hypotheses were not supported, this would not necessarily imply that the effect proposed did not actually exist, but rather that it was possible that the experiment had not adopted the appropriate design to detect such an effect. As such, the following hypotheses were tested:

1. **Non-solfège and taught solfège subjects will be more likely to yield lower Letter span scores than solfège subjects.**

2. **Non-solfège and taught solfège subjects will be more likely to yield lower Total task correct scores than solfège subjects.**

As in Experiment A (see Chapter 5.1), the taught solfège sample was extremely restricted in size, therefore statistical analyses were conducted using data from the solfège and non-solfège groups only; comparisons with the taught solfège group were made by reference to score frequencies and mean scores. It is also important to explain why taught solfège subjects were expected to reveal the same pattern of scores with non-solfège subjects in this experiment; the reasons for this grouping in the hypotheses were twofold:

1. As described in section 5.0, the last systematic contact taught solfège subjects had with solfège note names had taken place 9 months before the experiment. Although subjects participating in the experiment were confident about remembering and using solfège note names, it was considered highly unlikely that they would be inclined to use them spontaneously.

2. The aforementioned problem of a relative ambiguity in the distraction task data, owing to the mode of presentation of the task, became all the more relevant in the taught solfège group: even if subjects were inclined to use any form of note names in order to perform the task, the increased familiarity with letter names compared to solfège names, as well as the big time gap between the solfège lessons and the present experiment, was considered to diminish all possibility that taught solfège participant would use solfège note names without being prompted to by the experimental procedure itself, as was the case in Experiment A.

5.2.1.1. Method: Design

The study employed an independent design, using Solfège use as the independent variable, with three levels: solfège, taught solfège and non-solfège.

The dependent variable was performance on the memorisation task, measured by the following parameters:

1. Letter span, which was the total amount of letters remembered.

2. Total task correct, which was the total amount of correct answers given in the distraction task.

5.2.1.2. Method: Participants

46 musicians in total completed the task: 27 were English and 19 were Greek. 8 of the English participants had received solfège training over the previous academic year and were therefore familiar with solfège note names; the remaining 19 English participants had never received formal training in solfège.

There were 18 men and 9 women in the English group, with an average age of 23.48 years; 2 participants in the group were members of staff and the rest were students. 24 of the English participants reported they did not have AP and 3 that they did not know whether or not they had AP. There were 9 singers, 2 string players, 5 wind instrument players, 5 pianists, 5 guitar players and 1 percussionist in the English group.

The Greek group consisted of 8 men and 11 women, with an average age of 21.11 years. All Greek participants in this experiment were students and none reported to have AP. There were 3 singers, 1 string player, 2 wind instrument players, 11 pianists and 2 guitar players in the Greek group.

5.2.1.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were series of letters, with length varying between 2 and 7 items (e.g. K, P, M, U); the lower and upper thresholds of 2 and 7, respectively, were based on piloting. Letters A, B, C, D, E, F, G and H were excluded from the stimuli, as these letters are widely used to signify musical pitches in the letter notation system. All other letters were used; letter series were constructed in a way that:

1. Letters of the same sequence did not form acronyms, abbreviations or words.
2. No letter was repeated within the same sequence.
3. No letter appeared more than twice in the same experimental procedure.

Each letter of the sequence was presented in a separate screen, lasting 2 seconds; the presentation of the letters was only visual, with no accompanying sound. Letters were presented in a light grey background, in order to differentiate them from feedback and instruction screens, which were presented on white background. The decision to use background colour as a probe instead of a symbol, such as +, which is normally used in similar experiments, was made based on feedback from the pilot: the majority of students, most of them unfamiliar with psychology experiment routines, found the extra symbol less direct and more confusing than the colour change.

For the recall of the sequence, a screen appeared, containing a list of 16 options in total, featuring 9 letters and 7 'Blank' options. The 9 letters included all letters featuring in

the to-be-recalled sequence plus some irrelevant letters; irrelevant letters were controlled so that:

1. They did not contain letters from the immediately previous to-be-recalled sequence.

2. Like the to-be memorised sequences, they did not contain any of the first 8 letters of the English alphabet.

All options were presented with a green box next to them: subjects were instructed to choose the correct letters in the correct order by clicking on the relevant boxes; selected boxes turned red and acquired a serial number, indicating the order in which they had been clicked on. 'Blank' options were available in case a participant forgot or was dubious about a letter, but could remember its place in the sequence: in this event, they could click next to one of the 'Blank' options, in the respective order they thought that letter had occurred, and then proceed with the rest of the sequence.

A 'Clear' button was also provided, which could be used in case the subject made a mistake and needed to undo all previous selections for that sequence; there was no time restriction in the recall screen. Once the subject finished putting together their response, they clicked on 'Next' to proceed with the rest of the experiment. Before the commencement of the next task, a feedback screen appeared for 2 seconds, which informed subjects of the exact number of letters they had got correctly, in the correct order.

5.2.1.4. Materials and Procedure: Pitch comparison task stimuli

The pitch comparison task was identical to the one described in section 5.1.1.4, for Experiment A. Participants were presented with a series of note pairs, each pair consisting of a note in treble and a note in bass clef, in random order; for each pair, subjects were asked to indicate which note of the two was higher in pitch by pressing the relevant key on the numeric keypad. Each note pair stayed on screen for 4 seconds, during which subjects had to make their decision; if they failed to press any key during the 4 seconds, the experiment moved on to the next memorisation stimulus, counting the non-response as an incorrect answer. A feedback screen appeared for 1 second after

each pitch comparison task, informing participants whether or not they had got the answer right. The importance of performing the pitch comparison task as quickly and as accurately as possible was repeatedly stressed, both in the written and in the verbal instructions provided.

5.2.1.5. General procedure

As in Experiment A, participants had the chance to read extensive instructions and perform practice trials in order to become familiar with the procedure, before proceeding to the main experiment. The practice trials for Experiment B followed the same structure as in Experiment A: at first, subjects performed the memorisation task separately, then they performed the pitch comparison task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different sequence of letters was presented, which the subject was asked to memorise; before the presentation of each letter in the series, the pitch comparison task took place. There were between 2–7 distraction tasks in each experimental set and between 11–15 distraction tasks across the three experimental sets. A sample procedure of a single experimental set is shown in Figure 5.6:
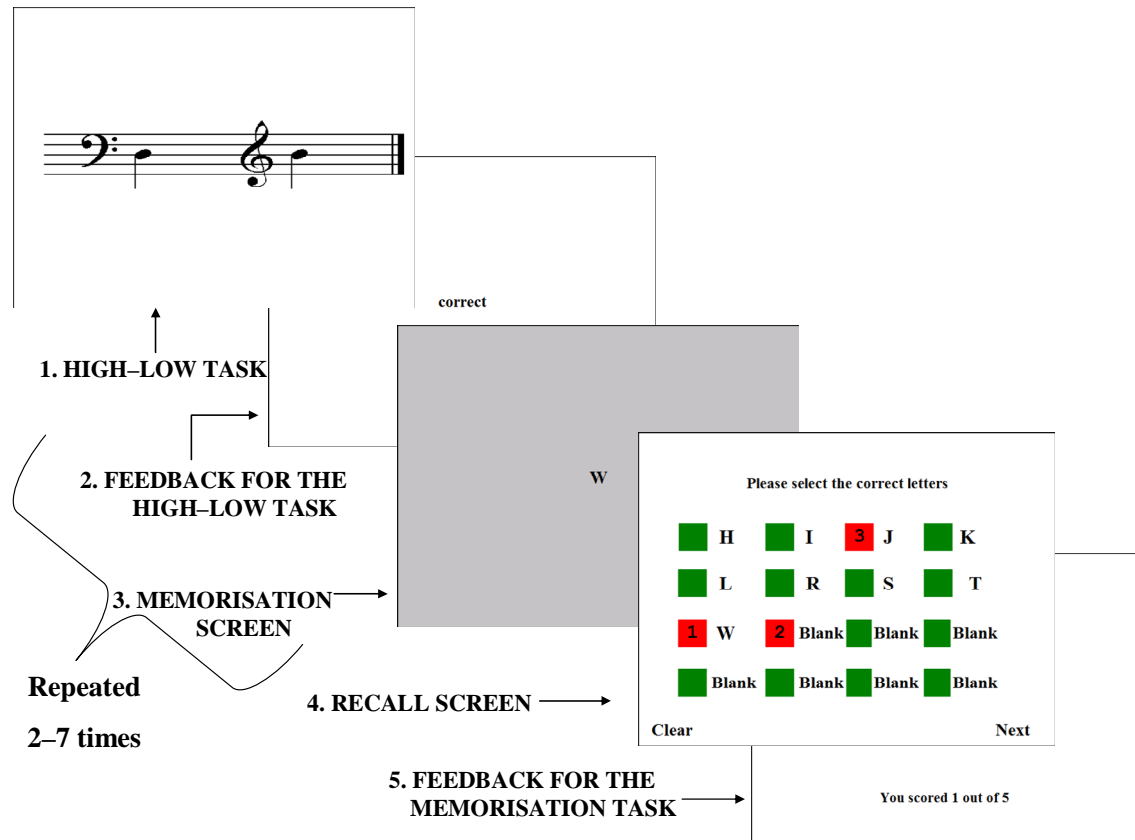
Figure 5.6: A sketch of the procedure followed in each of the three experimental sets in Experiment B1. Note that stages 1–3 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Letter span, which is equal to the total number of letters remembered correctly in the correct position, across all 3 experimental sets.

- Total errors, which is equal to the total number of errors made in the hi–low distraction task. This score is further broken down in two types of errors:

  - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

  - Accuracy errors, which is the total number of errors resulting from the subject answering incorrectly to the high–low task.

In addition to these scores, the number of *correct* answers given in the distraction task was also used in the analysis; it was calculated by subtracting the Total errors score

from the total number of high–low tasks performed throughout the 3 experimental sets. This was the **Total task correct** score and was calculated only as a percentage score.

5.2.1.6. Results

Table 5.5 displays the descriptive statistics for the memorisation and the distraction task scores separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	Median	SD	Min.	Max.	Skew	Kurtosis
Letter span	Solfège	10.21	10	2.97	5	15	- 0.17	- 0.70
	Taught solfège	12.88	14	2.85	7	15	- 1.41	1.78
	Non-solfège	11.16	13	3.61	3	15	- 0.65	- 0.45
Letter span (%)	Solfège	75.12	76.92	18.87	33.33	100	- 0.77	0.36
	Taught solfège	92.12	100	14.82	63.64	100	- 1.57	0.83
	Non-solfège	81.26	86.67	21.47	27.27	100	- 1	0.42
Total high–low errors	Solfège	1.95	2	1.55	0	6	0.91	1.08
	Taught solfège	2.75	2.50	2.12	0	7	1.03	1.80
	Non-solfège	1.95	1	2.09	0	7	1.29	0.87
Speed errors	Solfège	0.05	0	0.23	0	1	4.36	19
	Taught solfège	0.25	0	0.46	0	1	1.44	0
	Non-solfège	0.05	0	0.23	0	1	4.36	19
Accuracy errors	Solfège	1.89	2	1.52	0	6	1.04	1.50
	Taught solfège	2.50	2	2.20	0	7	1.28	1.85
	Non-solfège	1.89	1	2	0	7	1.28	1.07
Total task correct (%)	Solfège	85.86	86.67	11.16	53.85	100	- 1.15	2.50
	Taught solfège	79.54	80.91	16.21	46.15	100	- 1.16	2.35
	Non-solfège	85.54	90.91	14.60	53.33	100	- 1.01	0

Table 5.5: Descriptive statistics for B1 test scores, by solfège knowledge group. Note that group sizes were equal for the solfège and the non-solfège group (N = 19), but the taught solfège group was much smaller (N = 8).

Letter span and Total task correct scores were plotted, both collectively and separately for the solfège and non-solfège groups. As the number of letter stimuli was not constant across participants, percentage Letter span scores were used in order to perform meaningful comparisons; percentage Total task correct was also used for the same reason. In both cases, scores appeared to have significantly non-normal distributions, as shown in Table 5.6:

Tests of Normality

	Letter span (%)			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.925	.840	.868	.882	.858	.892
df	19	19	46	19	19	46
Sig.	.143	.005	.000	.023	.009	.000

Table 5.6: The Shapiro-Wilk test was preferred to the Kolmogorov-Smirnov because of the relatively small sample sizes. Test statistics are shown for the distribution of scores in the two groups separately as well as in the whole sample, which included solfège, taught solfège and non-solfège participants. All distributions were found significantly non-normal at $p < .05$, apart from the distribution of letter span in the solfège group.

Despite the violation of the normality assumption, MANOVA was conducted on the data and the Pillai-Bartlett trace was used, as it is considered robust to normality violations when sample sizes are equal. Using Pillai's trace, there was no significant effect of Solfège knowledge on Letter span and Total task correct scores, $V = 0.03$, $F_{(2, 35)} = 0.62$, $p > .05$, *ns*. As discussed in section 5.2.1, however, due to the design of the experiment, non-significant results should not be regarded as conclusive; consequently, further analyses were conducted on the data. Correlation analysis showed that Letter span scores were correlated with Total task correct scores in the total sample, $\tau = .40$, $p = .001$, indicating that subjects who performed well in the high–low task tended to yield high Letter span scores. Analysis in groups, however, revealed that this correlation did not exist in both groups: as in Experiment A, the two scores were uncorrelated in the solfège sample, $\tau = .31$, $p > .05$, *ns*, whereas a correlation was found in the non-solfège sample, $\tau = .46$, $p < .05$. Once more, correlation results revealed a strong tendency in

non-solfège participants to perform better in the letter task if they had also performed well in the intervening pitch task, whereas this tendency was not present amongst solfège participants. As the aim of the study was to compare results from participants who had performed both the distraction and the memorisation tasks as accurately as possible, excluding subjects who had favoured one task over the other, correlation analysis was conducted only for data from subjects whose Total task correct score was over 70%. The equivalent cut-off threshold in Unsworth, Heitz et al's Aospan tests was 85% (Unsworth et al, 2005); a lower threshold was chosen in the present study, based on sample size and sample means: as the mean Total task correct score was 85.70% and the median was just slightly higher, at 88.79%, eliminating scores below 85% would mean eliminating almost half the sample, reducing it to 21 subjects in total. Total task correct scores over 70%, on the other hand, were still considered to express a satisfactory performance level in the distraction task and represented 86.8% of the total sample, therefore the 70% threshold was preferred for this study.

The Kolomogorov-Smirnov Z test statistic was preferred as it is considered to have more power than the Mann-Whitney U in samples where the number of cases in each group is below 25 (Field, 2009); it was found non-significant at p > .05, indicating that, for subjects who had performed correctly in 70% or above of the distraction tasks, the distribution of Letter span scores was the same between solfège and non-solfège subjects. An examination of score frequencies and means revealed that, in fact, non-solfège participants had produced a higher mean Letter span score than solfège participants: the mean percentage of the total letters remembered correctly was 77.15% for solfège and 86% for non-solfège participants; 8 out of the 15 subjects in the non-solfège group had recalled correctly over 90% of the letter stimuli, whereas the ratio in the solfège group was 4 out of 18 subjects. Finally, correlation analysis showed that, in this high-span group, the positive relationship between percentage Letter span scores and Total task correct scores was still valid for the non-solfège group, $\tau = .45$, $p < .05$, whereas it did not appear in the solfège group, $\tau = .24$, $p > .05$.

Regarding Letter span, the taught solfège group (N = 8) yielded a pattern of performances closer to the non-solfège rather than the solfège group: the mean percentage of letters taught solfège subjects had remembered correctly was 92.12%, with 6 out of the 8 subjects recalling over 90% of the letter stimuli; it is important to

note that only one of these 8 participants had scored below the 70% threshold in the Total task correct score.

In terms of the pitch distraction task, both the Mann-Whitney and the Kolmogorov-Smirnov test statistics were found non-significant at $p > .05$, suggesting that the distribution of Total task correct scores was the same in the solfège and the non-solfège group. Table 5.7 shows the mean Total task correct scores for the taught solfège, solfège and non-solfège groups, with their associated standard errors:

Total task correct (%) scores

	Mean	**SE**
Solfège	85.86	2.56
Non-solfège	85.54	3.35
Taught solfège	79.54	5.73

Table 5.7: Mean Total task correct scores for the three solfège groups.

As shown in Table 5.7, participants in the three groups displayed a similar pattern of performances to the one in Experiment A: the mean Total task correct score for the taught solfège group was lower compared to the respective mean of the other two groups. Score frequencies were checked and revealed that, contrary to the results in Experiment A, the difference in means was caused by an outlier: a subject in the taught solfège group who had achieved a Total task correct score of 46.1%. Although this was the lowest performance across groups, comparable values could be found in the solfège and non-solfège group, where the lowest Total task correct performances were equal to 53.85% and 53.33%, respectively. Consequently, the difference in means between the taught solfège and the other two groups was attributed to the small number of participants in the former, which caused lower scores to have a bigger impact on the mean score.

5.2.1.7. Summary

Hypothesis no. 1 stated that:

Non-solfège and taught solfège subjects will be more likely to yield lower Letter span scores than solfège subjects.

This hypothesis was not supported by data analysis; non-parametric test statistics showed that there was no significant effect of solfège knowledge on Letter span; moreover, Letter span score means were lower for the solfège group than for the other two groups. These findings were replicated in the condition in which only participants scoring over the 70% threshold in the distraction task were used, indicating that the non-significant effect of solfège knowledge was valid and did not depend on the level of engagement with the two parallel tasks. Although, as mentioned in section 5.2.1, the experimental design did not guarantee that all subjects were performing the distraction task using the same mechanisms, this primary evidence suggested that the knowledge of letter names, as opposed to solfège names, did not constitute an inherent drawback in the parallel performance of a phonemic and a music reading task, as expected.

Hypothesis no. 2 stated that:

Non-solfège and taught solfège subjects will be more likely to yield lower Total task correct scores than solfège subjects.

Data analysis, however, did not reveal any statistical significance in the difference between solfège and non-solfège participants in the distraction task; mean percentage scores for the total amount of correct answers in the task were similar across all three groups, solfège, non-solfège and taught solfège. These findings did not support Hypothesis no. 2, which was based on the premise that the familiarity with letter names, as opposed to solfège names, when combined with a letter-based task, would cause a conflict as the performance of both tasks would require sharing of the same cognitive resources.

Letter span and Total task correct scores were found to be significantly correlated in data for non-solfège participants, whereas this correlation was not found significant in data from the solfège group; these relationships were valid both in data from the whole

sample and in data from participants who had performed above the 70% threshold in the distraction task. The same effect had been found in Experiment A, suggesting a pattern in the non-solfège group, whereby the most efficient memorisers were also most competent in the high–low task. As this effect was not traced in the solfège group it was not possible to draw definite conclusions; data from these two experiments as well as from subsequent experiments regarding this observed difference between solfège and non-solfège subjects will be examined and interpreted in section 5.4.

5.2.2. Experiment B2: Musical interference in a morphemic memorisation task

For Experiment B2, participants were asked to memorise nonsense syllable sequences while simultaneously performing a visual pitch discrimination task, identical to the pitch discrimination task in Experiment A (Chapter 5.1.1.4). Nonsense syllables were used in this experiment in order to compare results with the subsequent experiment, B3, where subjects were asked to memorise monosyllabic words; if an effect of solfège knowledge was to be found in either condition, it would provide a valuable contribution towards understanding the nature of solfège syllables within memory and cognition.

For the nonsense syllable condition, as well as for the monosyllabic word condition in the following experiment, solfège knowledge was expected to have a detrimental effect on syllable recall: the assumed use of solfège syllables in the distraction task was expected to interfere with the recall of syllables in the memorisation task. As in Experiment B1, a major restriction in the design of the experiment was that the mode of performance of the distraction task was not specified; therefore, the interpretation of results could not be conclusive without further support from more extensive research. For this study the following hypotheses were tested:

1. **Solfège subjects will be more likely to yield lower Syllable span scores than non-solfège and taught solfège subjects.**

2. **The distribution of correct answers in the distraction task will be similar across the three groups.**

Taught solfège subjects were grouped together with non-solfège subjects in this experiment for the same reasons described in section 5.2.1; the non-restrictive presentation of the distraction task in the whole of Experiment B was considered to decrease the possibility of taught solfège participants actually using solfège note names to perform the task.

5.2.2.1. Method: Design

The design of the experiment was identical with the design of B1: the experiment employed an independent design, using Solfège use as the independent variable, with three levels: solfège, taught solfège and non-solfège.

The dependent variable was performance on the memorisation task, measured by the following parameters:

3. Syllable span, which was the total amount of syllables remembered.

4. Total task correct, which was the total amount of correct answers given in the distraction task.

5.2.2.2. Method: Participants

A total of 44 musicians took part in the experiment: 24 were English and 20 were Greek. From the English participants, 5 had received solfège training over the previous academic year and were familiar with solfège note names; the remaining 19 English participants were unfamiliar with solfège.

The English group included 15 men and 9 women, with an average age of 23.88 years; 2 subjects were members of staff and the rest were students. 1 person in the group was an AP possessor, 3 had answered they did not whether or not they possessed AP and the remaining participants did not possess AP. There were 8 singers, 2 string players, 5 wind instrument players, 5 pianists, 3 guitar players and 1 percussionist in the group.

The Greek group included 6 men and 14 women, with and average age of 21.35 years; the group consisted entirely of students. There were 2 AP possessors in the group and 1 person who had reported she did not know whether or not she had AP. 2 subjects were first study singers, 1 was a string player, 1 a wind player, 14 were pianists and 1 a percussionist.

5.2.2.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were series of nonsense syllables (e.g. ul, qo, ik) with length varying between 2 and 7 items. Nonsense syllables were constructed so that:

1. Each one comprised one consonant and one vowel, in any order.
2. The two letters forming the syllable did not form an acronym or abbreviation.

Nonsense syllable sequences were also constructed so that they did not form acronyms, abbreviations or words within them. Each syllable of the sequence was presented in a separate screen, lasting 2 seconds; the presentation of the syllables was visual only, with no accompanying sound. Each syllable was presented in a light grey background, in order to avoid confusion between memorisation stimuli, instructions and feedback screens.

For the recall of the sequence, a screen appeared, containing a list of 16 options in total, featuring 9 nonsense syllables and 7 'Blank' options. The 9 syllables included all letters featuring in the to-be-recalled sequence plus some irrelevant syllables; irrelevant syllables were controlled so that:

1. They did not contain syllables from the immediately previous to-be-recalled sequence.
2. They fulfilled the same aforementioned criteria as the to-be-recalled syllables.

All options were presented with a green box next to them: subjects were instructed to choose the correct syllables in the correct order by clicking on the relevant boxes; selected boxes turned red and acquired a serial number, indicating the order in which they had been clicked on. 'Blank' options were provided in case a participant forgot or was dubious about a syllable, but could remember its place in the sequence.

A 'Clear' button was also provided, which could be used in case the subject made a mistake and needed to undo all previous selections in that screen; there was no time restriction in the recall screen. Once the subject finished putting together their response, they clicked on 'Next' to proceed with the rest of the experiment. Before the commencement of the next task, a feedback screen appeared for 2 seconds, which

informed subjects of the exact number of letters they had got correctly, in the correct order.

5.2.2.4. Materials and Procedure: Pitch comparison task stimuli

The pitch comparison task was identical to the one described in section 5.1.1.3, for Experiment A. Participants were presented with a series of note pairs, each pair consisting of a note in treble and a note in bass clef, in random order; for each pair, subjects had 4 seconds to indicate which note of the two was higher in pitch by pressing the relevant key on the numeric keypad. A feedback screen appeared for 1 second after each pitch comparison task, informing participants whether or not they had got the answer right. The importance of performing the pitch comparison task as quickly and as accurately as possible was repeatedly stressed, both in written and in verbal instructions.

5.2.2.5. General procedure

As in Experiment B1, participants had the chance to read extensive instructions and perform practice trials in order to get familiarised with the procedure, before proceeding to the main experiment. The practice trials for Experiment B2 followed the same structure as in Experiment B1: at first, subjects performed the memorisation task separately, then they performed the pitch comparison task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different sequence of syllables was presented, which the subject was asked to memorise; before the presentation of each syllable in the series, the pitch comparison task took place. The number of distraction tasks performed in each experimental set varied between 2–7; the number of distraction tasks performed during the three experimental sets varied between 11–15. A sample procedure of a single experimental set is shown in Figure 5.7:
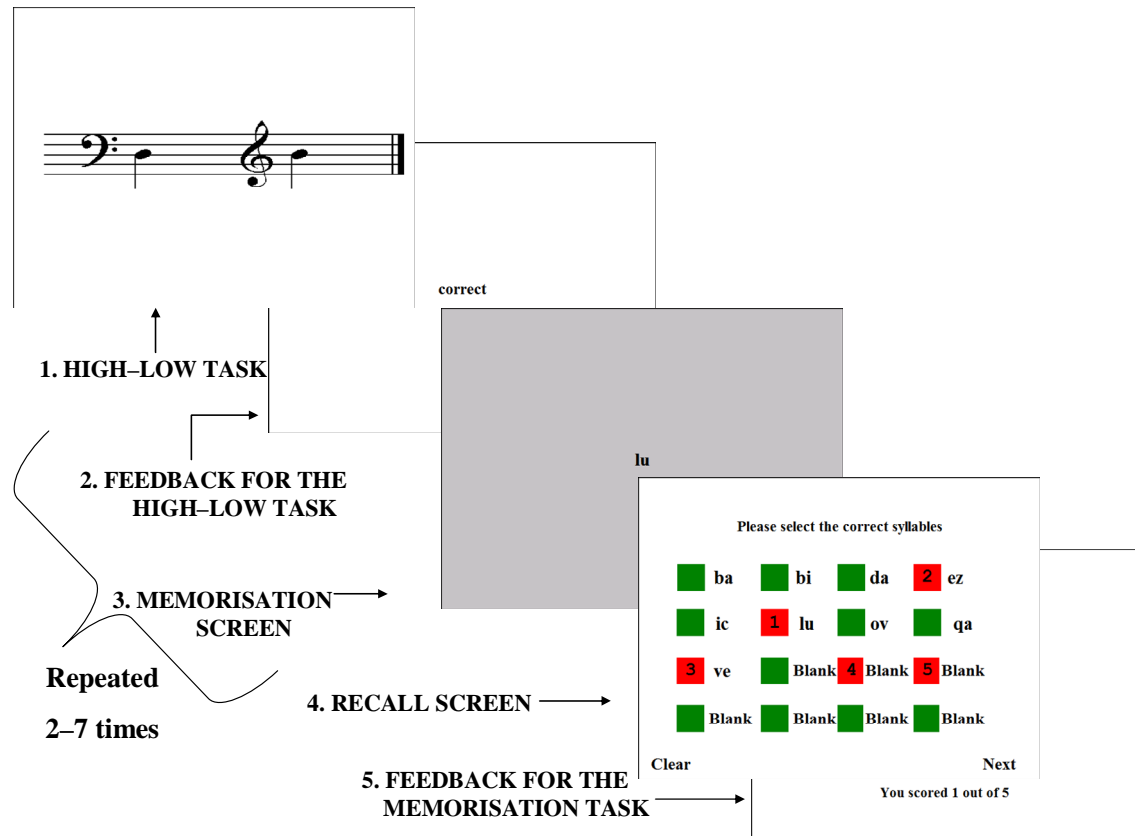
Figure 5.7: A sketch of the procedure followed in each of the three experimental sets in Experiment B2. Note that stages 1–3 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Syllable span, which is equal to the total number of syllables remembered correctly in the correct position, across all 3 experimental sets.

- Total errors, which is equal to the total number of errors made in the hi–low distraction task. This score is further broken down in two types of errors:

  - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

  - Accuracy errors, which is the total number of errors resulting from the subject answering incorrectly to the high–low task.

In addition to these scores, the number of correct answers given in the distraction task was also used in the analysis; this was calculated by subtracting the Total errors score

from the total number of high–low tasks performed throughout the 3 experimental sets. This was the **Total task correct** score and was calculated only as a percentage score.

5.2.2.6. Results

Table 5.8 displays the descriptive statistics for the memorisation and the distraction task scores separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	**Median**	**SD**	**Min.**	**Max.**	**Skew**	**Kurtosis**
Syllable span	Solfège	13.65	15	1.63	11	15	- 0.58	- 1.42
	Taught solfège	14.20	15	1.79	11	15	- 2.23	5
	Non-solfège	14	15	1.41	11	15	- 0.92	- 0.78
Syllable span (%)	Solfège	59.55	60	20.01	20	100	0.03	- 0.20
	Taught solfège	80.85	73.33	13.98	66.67	100	0.66	- 1.74
	Non-solfège	76.53	80	23.33	20	100	- 1	0.54
Total high–low errors	Solfège	2.50	2	2.01	0	8	1.03	1.48
	Taught solfège	2.60	2	2.19	0	6	0.85	1.75
	Non-solfège	1.89	2	1.56	0	6	1.08	1.26
Speed errors	Solfège	0.30	0	0.47	0	1	0.95	- 1.24
	Taught solfège	0	0	0	0	0		
	Non-solfège	0.11	0	0.32	0	1	2.80	6.51
Accuracy errors	Solfège	2.20	2	1.91	0	8	1.35	3.31
	Taught solfège	2.60	2	2.19	0	6	0.85	1.75
	Non-solfège	1.79	2	1.55	0	6	1.09	1.66
Total task correct (%)	Solfège	81.88	85.64	14.27	46.67	100	- 0.80	0.34
	Taught solfège	81.21	86.67	15.28	60	100	- 0.37	- 0.35
	Non-solfège	86.42	86.67	11.88	50	100	- 1.64	3.95

Table 5.8: Descriptive statistics for B2 test scores, divided by solfège knowledge group. Note the difference in group sizes: for the solfège group N = 20, for the taught solfège group N = 5 and for the non-solfège group N = 19.

Syllable span and Total task correct scores were plotted, both collectively and separately for the solfège and non-solfège groups. As the number of syllable stimuli was not constant across participants, percentage span scores were used in order to perform meaningful comparisons. Both variables appeared to have significantly non-normal distributions of scores for the non-solfège group, as shown in Table 5.9:

Tests of Normality

	Syllable span (%)			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.990	.869	.945	.939	.854	.909
df	20	19	44	20	19	44
Sig.	.998	.014	.037	.226	.008	.002

Table 5.9: The Shapiro-Wilk test was chosen because of the relatively small sample sizes. Test statistics are shown for the distribution of scores in the two groups separately as well as in the whole sample, including solfège, taught solfège and non-solfège participants. Distributions of both scores were found significantly non-normal at $p < .05$ for the non-solfège group, whereas the solfège group produced normally distributed scores ($p > .05$).

Because of the non-normal distribution of scores in the non-solfège group, non-parametric tests were selected for the analysis. The Kolmogorov-Smirnov Z test was conducted in the Syllable span scores for the solfège and non-solfège groups and was found significant at $p < .05$, indicating that the distribution of scores was not the same across groups. Simple regression analysis was conducted in order to determine the nature of the effect, using Syllable span scores as the outcome variable and solfège knowledge as the predictor variable; as the residuals plots raised concerns regarding heteroscedasticity in the data, the analysis was based on bootstrap samples. Results from the analysis are shown in Table 5.10:

Syllable span scores

	B	SE B[a]	β
Constant	76.53	5.31	
Solfège knowledge	- 16.98	6.68	- .37**

$R^2 = .14$, p $< .05$

** p $< .05$, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap samples.

Table 5.10: Simple regression results for Syllable span scores.

The value of R^2 indicates that solfège knowledge as a predictor affected the model significantly; the value of β suggested that solfège subjects had yielded significantly lower Syllable span scores than non-solfège subjects. Although this supported the first experimental hypothesis, further validation was sought by conducting the same regression analysis in the data of participants who had scored over 70% correctly in the high–low distraction task; for an explanation regarding the selection of the 70% threshold, see section 5.2.1.6. Out of the 39 solfège and non-solfège participants, 5 had scored below 70% correctly in the distraction task; data from the remaining 34 subjects, 16 solfège and 18 non-solfège, were used in the analysis; results are shown in Table 5.11:

Syllable span scores

	B	SE B[a]	β
Constant	75.23	– 0.30	
Solfège knowledge	– 11.99	– 0.11	– .28*

$R^2 = .08$, p $> .05$

* p $> .05$, n*s*

a. Bootstrapped values. Significance for β is also based on bootstrap samples.

Table 5.11: Simple regression results for Syllable span scores for data from participants with over 70% correct answers in the high–low task.

For participants with Total task correct scores over 70, solfège knowledge did not appear to have a significant effect on Syllable span scores; although the relationship between solfège group and Syllable span appeared to have the same direction as in the previous results, with solfège subjects yielding lower Syllable span scores, this effect was not significant in the sample, including only participants who were over the 70% threshold in the distraction task. The contradiction between the two regression analysis results called for a closer examination of the scores of the 5 excluded participants, 4 of whom belonged in the solfège and 1 in the non-solfège group. 3 out of the 4 solfège participants who had scored below 70% correctly in the high–low task had also yielded Syllable span scores that were much lower than the solfège group mean; on the other hand, the non-solfège subject who had not crossed the 70% threshold in the distraction task had a Syllable span of 100%. These findings partially illuminated the contradiction: the excluded solfège subjects had lowered the group mean considerably, whereas the excluded non-solfège subject had contributed in raising the respective group mean; the combination of these opposing trends in the opposite groups had produced a magnified effect of comparatively poorer performance in the solfège group.

Although the sample in which the initial regression analysis was conducted was relatively small (N = 39), the significant skew in results caused by just 5 participants deserved further investigation: the overall relationship between Syllable span and Total task correct scores was examined, in order to determine whether the 4 solfège subjects were part of a generalized trend to improve or deteriorate their performance in the Syllable span and Total task correct scores in parallel; such a finding would be consistent with results from Experiments A and B2 (see sections 5.1.2 and 5.2.1.6). Correlation analysis, however, showed Syllable span to be uncorrelated with Total task correct score in the overall sample, $\tau = .18$, $p > .05$, *ns*, as well as within subgroups: $\tau = .18$, $p > .05$, *ns*, for the solfège group and $\tau = .10$, $p > .05$, *ns*, for the non-solfège group. The correlation between Syllable span and Total task correct was also found non-significant in the filtered sample, containing only participants who had surpassed the 70% threshold.[1]

---

[1] For the sample including participants who had scored over 70% correctly in the distraction task, correlation results between Syllable span and Total task correct scores were the following: $\tau = .14$, $p > .05$, *ns*, for the overall sample, $\tau = .04$, $p > .05$, *ns*, for the solfège group and $\tau = .21$, $p > .05$, *ns*, for the non-solfège group.

Regarding Syllable span scores across all three groups, the scores for the 5 subjects in the taught solfège group are displayed in Figure 5.8:

Taught solfege group
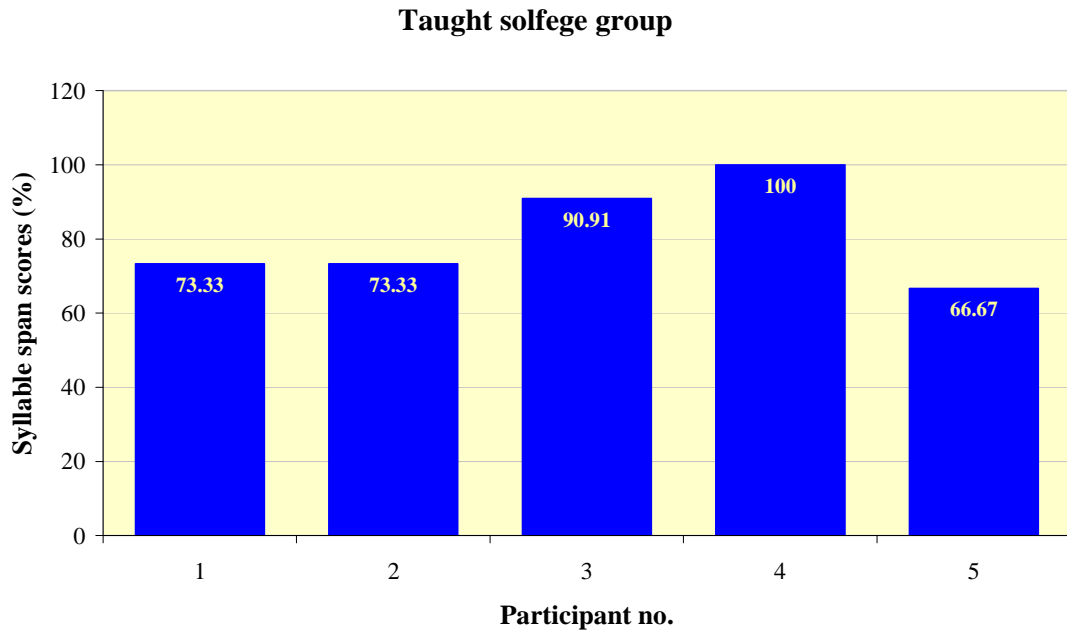
Figure 5.8: Taught solfège subjects' data for Syllable span. The values in bars represent percentage scores.

All 5 taught solfège subjects yielded relatively high Syllable span scores; the lowest score in the taught solfège group was 66.67. The minimum scores in the solfège and non-solfège groups, for the filtered sample of participants who had scored over 70% in the distraction task, were 33.33 for the solfège and 20 for the non-solfège group, respectively. It is important to note at this point that participant no. 2 from the taught solfège group had scored below 70% correctly in the distraction task; with the exclusion of this subject's score from the comparison, the difference in mean Syllable span scores for the three groups grew even greater, as shown in Figure 5.9:

Mean Syllable span scores (%)
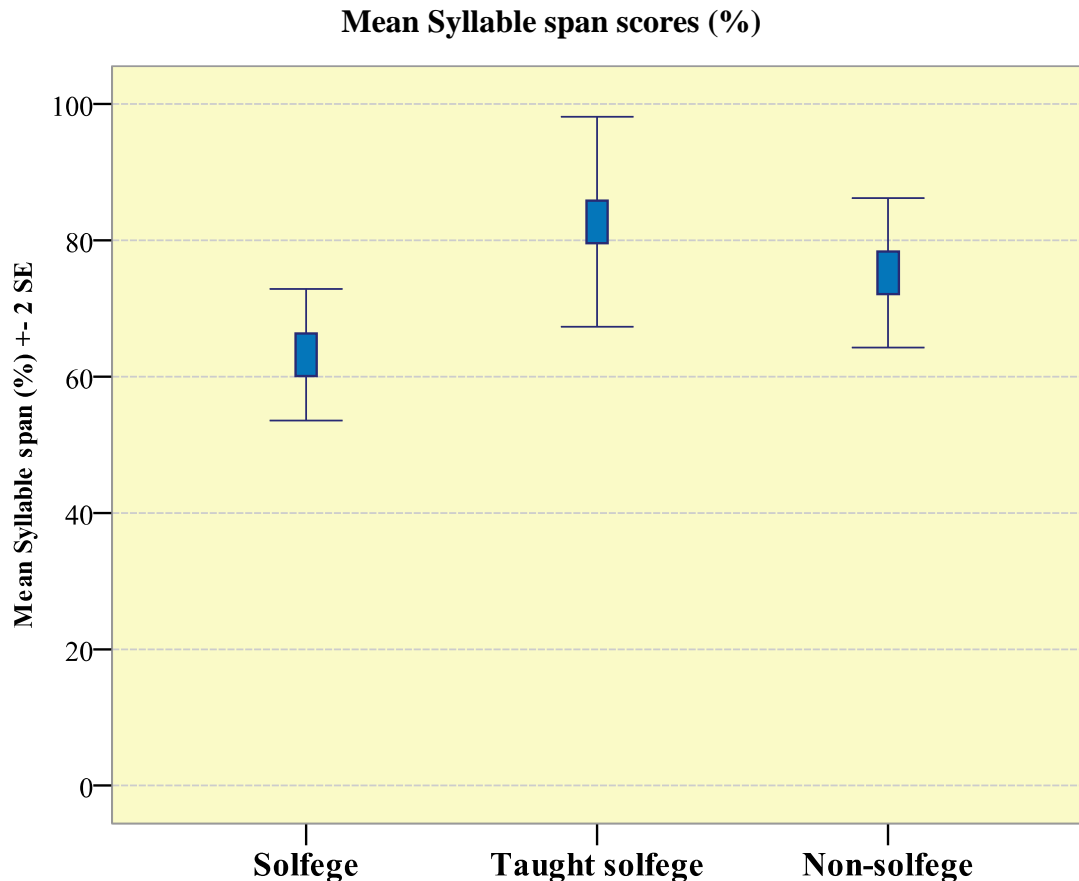
Figure 5.9: Mean Syllable span scores per solfège knowledge group. Mean scores for all three groups were calculated including only participants who had scored over 70% correctly in the distraction task; for the taught solfège group (N = 4) M = 82.73, for the solfège group (N = 16) M = 63.24 and for the non-solfège group (N = 18) M = 75.23. Error bars represent the mean ± 2 standard errors.

As shown in Figure 5.9, taught solfège participants yielded the highest percentage scores in the syllable memorisation task; the difference between mean scores, however, was not evaluated further or interpreted in the analysis, because of the small number of participants in the taught solfège group.

Regarding performance scores in the distraction task, the Kolmogorov-Smirnov test was conducted on the Total task correct data and the test statistic was found non-significant at p > .05, indicating that the distribution of scores was similar for solfège and non-solfège participants. Mean Total task correct scores with their associated standard errors are presented in Table 5.12:

Total task correct (%) scores

	Mean	SE
Solfège	81.88	3.19
Non-solfège	86.42	2.73
Taught solfège	81.21	6.84

Table 5.12: Mean Total task correct scores for the three groups in Experiment B2.

Mean Total task correct scores displayed in Table 5.13 suggest that the performance in the distraction task was very similar in the taught and the other two groups; the non-solfège group showed a slightly elevated mean, but, as mentioned previously, this difference was not found statistically significant.

5.2.2.7. Summary

Hypothesis no. 1 stated that:

> Solfège subjects will be more likely to yield lower Syllable span scores than non-solfège and taught solfège subjects.

This hypothesis was not supported conclusively by analysis; although initial non-parametric analysis yielded significant results, indicating that solfège participants had produced significantly lower scores in the syllable memorisation task, further filtering of participants cancelled this effect, after 5 participants who had performed below 70% correctly in the distraction task were removed from the analysis. The decision to exclude of participants who had answered less than 70% correctly in the distraction task had been established in Experiment B1 (section 5.2.1.6), in order to filter out participants who were neglecting performance in the distraction task, presumably in an attempt to perform better in the memorisation task. Kendall's statistic in the present experiment, however, revealed that there was no correlation, either positive or negative, between Syllable span and Total task correct scores; this suggested that participants who performed poorly in the distraction task were not, in fact, favouring one task over

the other, but were rather facing genuine difficulties in completing the tasks accurately. A closer examination of the data from participants who had been excluded due to failure to cross the 70% threshold in the distraction task revealed that, out of the 5 participants, only one had actually yielded a high score in the memorisation task; the remaining 4 participants had low scores in both tasks, 3 of them producing scores below the group mean in Syllable span. Although these findings suggested that scores lower than 70 in the Total task correct did not, in most cases, represent a failure to complete the experiment efficiently but rather a restriction in the ability to perform both tasks simultaneously, the existence of one participant who had performed poorly in the distraction task and 100% correctly in the memorisation tasks was considered to justify the preservation of the 70% threshold as a 'safety net' against subjects who might have failed to follow instructions and pay equal attention to performing both tasks as efficiently as possible. The ambiguity, however, was taken into consideration in the general discussion in Chapter 5.4; testing a much larger sample would certainly help to clarify whether or not solfège subjects were likely to perform worse in the syllable memorisation task.

Hypothesis no. 2 stated that:

> Solfège subjects will be more likely to yield lower Total task correct scores than non-solfège and taught solfège subjects.

This hypothesis was not supported by data analysis; non-parametric tests conducted on the solfège and the non-solfège data did not reveal a significant effect of group in the scores. Taught solfège subjects were not included in statistical analyses because of their extremely limited numbers; observation of score frequencies and mean Total task correct scores, however, revealed that it was likely that they had produced a similar pattern of results with the other two groups.

5.2.3. Experiment B3: Musical interference in a lexical memorisation task

Participants in Experiment B3 were asked to memorise monosyllabic word sequences while simultaneously performing a visual pitch discrimination task, identical to the pitch discrimination task in Experiment A (section 5.1.1.4). Solfège knowledge was expected to have a detrimental effect on word recall, as the presumed use of solfège names in the distraction task was expected to interfere with the recall of lexical items in the memorisation task. The following hypotheses were tested:

1. **Solfège participants will be more likely to yield lower Word span scores compared to the non-solfège and the taught solfège group.**

2. **Solfège participants will be more likely to yield lower Total task correct scores compared to non-solfège and taught solfège participants.**

As in Experiments B1 and B2, the mode of performance in the distraction task was not specified, therefore no conclusive evidence could be extracted from data analysis.

5.2.3.1. Method: Design

B3 followed the independent design of the other experiments in group B: Solfège use was the independent variable, having three levels: solfège, taught solfège and non-solfège; performance on the memorisation task was the dependent variable, measured by:

1. Word span, which was the total amount of words remembered.
2. Total task correct, which was the total amount of correct answers given in the distraction task.

5.2.3.2. Method: Participants

A total of 46 subjects took part in the experiment: 20 were Greek and 26 English. Out of the 26 English participants, 7 had been taught solfège during the previous academic year. A total of 18 men and 8 women participated in the English group; their mean age was 23.38 years. 3 of the English participants were members of staff, while the remaining were students; there were no AP possessors in the group, although one subject had responded he did not know whether or not he had AP. In terms of instrumental categories, there were 12 singers, 3 wind players, 5 pianists, 4 guitar players and 2 percussionists.

The Greek group included 4 men and 16 women, with a mean age of 22.70 years. There was only one staff member in the Greek group and one person who reported she did not know whether she was an AP possessor; all the other subjects were non-AP students. There was 1 string player, 1 wind player, 15 pianists and 1 guitar player in the group.

5.2.3.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were series of monosyllabic words (e.g. the, as, pour), with the length of each sequence varying between 2 and 7 items. Words in a single sequence were controlled so that they did not construct meaningful sentences. Each word of the sequence was presented in a separate screen, lasting 2 seconds; the presentation was visual only, with no accompanying sound. Each word was presented in a light grey background, in order to avoid confusion between memorisation stimuli, instructions and feedback screens.

For the recall of the sequence, a screen appeared, containing a list of 16 options in total, featuring 9 words, including all the words in the to-be-recalled sequence, and 7 'Blank' options. As in Experiments B1 and B2, each option had a green box next to it, which the subject had to click on in order to activate his/her choice; participants had to indicate the correct words in the correct order. The 'Clear' button could be used in order to undo all

selections in the screen and the 'Next' button led to the feedback screen of the memorisation task and, subsequently, to the rest of the experiment.

5.2.3.4. Materials and Procedure: Pitch comparison task stimuli

The pitch comparison task was identical to the previous experiments. Participants were presented with a series of note pairs, each pair consisting of a note in treble and a note in bass clef, in random order; for each pair, subjects had 4 seconds to indicate which note of the two was higher in pitch by pressing the relevant key on the numeric keypad. A feedback screen appeared for 1 second after each task; the importance of performing the pitch comparison task as quickly and as accurately as possible was repeatedly stressed, both in written and in verbal instructions.

5.2.3.5. General procedure

As in Experiments B1 and B2, participants had the chance to read extensive instructions and perform practice trials in order to get familiarised with the procedure, before proceeding to the main experiment. The practice trials for Experiment B3 followed the same structure as in Experiments B1 and B2: at first, subjects performed the memorisation task separately, then they performed the pitch comparison task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different sequence of words was presented, which the subject was asked to memorise; before the presentation of each word in the series, the pitch comparison task took place. The number of distraction tasks performed in each experimental set was between 2–7; the number of distraction tasks performed during the three experimental sets was between 11–15. A sample procedure of a single experimental set is shown in Figure 5.10:
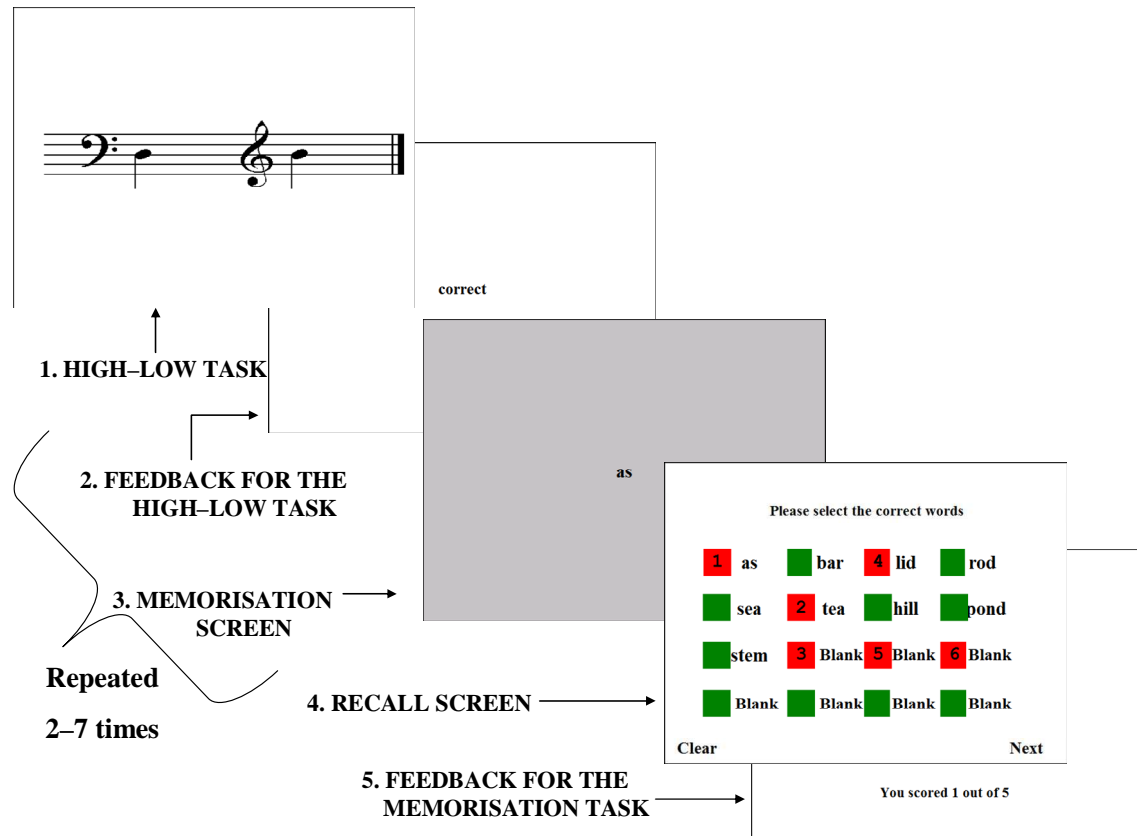
Figure 5.10: A sketch of the procedure followed in each of the three experimental sets in Experiment B3. Note that stages 1–3 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Word span, equal to the total number of words remembered correctly in the correct position, across all 3 experimental sets.
- Total errors, which is equal to the total number of errors made in the high–low distraction task. This score was further broken down in two types of errors:
  - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.
  - Accuracy errors, which is the total number of errors resulting from the subject answering incorrectly to the high–low task.

In addition to these scores, the number of correct answers given in the distraction task was also used in the analysis; this was calculated by subtracting the Total errors score

from the total number of high–low tasks performed throughout the 3 experimental sets. This was the **Total task correct** score and was calculated only as a percentage score.

5.2.3.6. Results

Table 5.13 displays the descriptive statistics for the memorisation and the distraction task scores separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	Median	SD	Min.	Max.	Skew	Kurtosis
Word span	Solfège	10	11	2.99	3	15	- 0.96	0.94
	Taught solfège	12.29	12.43	2.69	7	15	- 1.43	2.22
	Non-solfège	11.68	12	2.34	6	15	- 0.89	0.42
Word span (%)	Solfège	74.34	74.17	21.82	20	100	- 1.07	0.89
	Taught solfège	89.09	93.33	14.66	63.64	100	- 1.21	- 0.10
	Non-solfège	86.12	93.33	15.04	54.55	100	- 0.76	- 0.79
Total high–low errors	Solfège	2.25	2	2.07	0	7	0.97	0.12
	Taught solfège	2.71	2	2.56	0	8	1.69	3.72
	Non-solfège	2.05	2	1.81	0	6	1.11	0.59
Speed errors	Solfège	0.10	0	0.31	0	1	2.89	7.04
	Taught solfège	0.14	0	0.38	0	1	2.65	7
	Non-solfège	0.11	0	0.32	0	1	2.80	6.51
Accuracy errors	Solfège	2.15	1.50	2.03	0	7	1.07	0.47
	Taught solfège	2.57	2	2.57	0	8	1.90	4.39
	Non-solfège	1.95	1	1.87	0	6	1.11	0.48
Total task correct (%)	Solfège	82.72	86.67	16.08	45.46	100	- 1.01	0.24
	Taught solfège	80.57	84.61	17.35	46.67	100	- 1.33	2.31
	Non-solfège	84.86	86.67	13.20	53.85	100	- 1.07	0.48

Table 5.13: Descriptive statistics for B3 test scores, divided by solfège knowledge group. For the solfège group N = 20, for the taught solfège group N = 7 and for the non-solfège group N = 19.

Word span and Total task correct scores were plotted, both collectively and separately for the solfège and non-solfège groups. As the number of word stimuli was not constant across participants, percentage span scores rather than absolute values were used for the analyses. In all cases, both variables appeared to have significantly non-normal distributions, as shown in Table 5.14:

Tests of Normality

	Word span (%)			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.891	.836	.864	.890	.881	.883
df	20	19	46	20	19	46
Sig.	.028	.004	.000	.027	.022	.000

Table 5.14: The Shapiro-Wilk test was chosen because of the relatively small sample sizes. Test statistics are shown for the distribution of scores in the two groups separately as well as in the whole sample, including solfège, taught solfège and non-solfège participants. Distributions of both scores were found significantly non-normal at $p < .05$ for all groups.

Because of the non-normal distribution of scores in the non-solfège group, non-parametric tests were conducted on the data: the Kolmogorov-Smirnov test statistic was found non-significant at $p < .05$, indicating that the distribution of Word span scores was the same in the solfège and the non-solfège groups. The same test was conducted on data from participants who had crossed the 70% correct threshold in the distraction task and was found again non-significant at $p > .05$. It is worth noting, however, that when simple regression analysis was conducted using Word span as the outcome variable and solfège group as the predictor, the significance of F change was $p = .058$, which was close to the significance threshold; this borderline non-significance of a negative correlation, in combination with the methodological restriction of the experiment discussed in section 5.2.1., indicated that testing a bigger sample and perhaps modifying the presentation of the distraction task was necessary in order to draw definite conclusions regarding the effect of solfège knowledge on Word span.

Taught solfège, solfège and non-solfège mean Word span scores were compared both for the total sample and for the filtered sample, in which participants had achieved over 70% correctly in the high–low task; scores are shown in Figure 5.11:

Mean Word span scores (%)
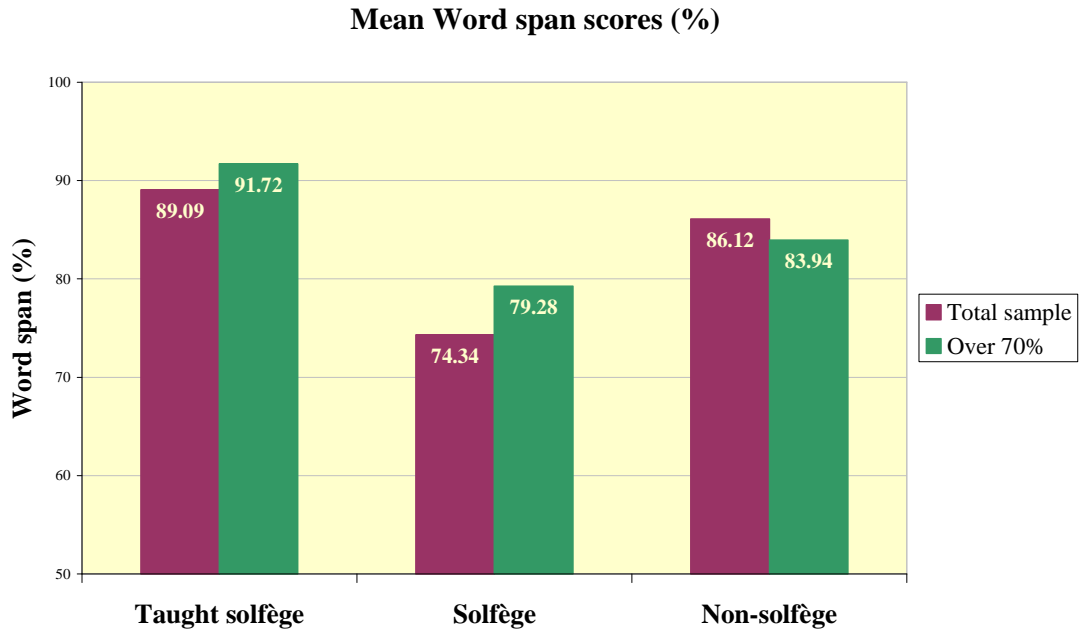
Figure 5.11: Mean Word span scores in each solfège group. Purple bars represent data for groups including all participants and green bars represent data from groups excluding participants who had scored lower than 70% correctly in the distraction task; percentage Word span scores are displayed inside the bar of each group.

As shown in Figure 5.11, taught solfège subjects had the highest mean Word span scores and solfège subjects had the lowest. The difference in mean scores between the solfège and the non-solfège groups was higher in the total sample than in the filtered sample, which included only participants scoring 70% or above in the distraction task; this indicated that solfège subjects scoring below 70% had mostly yielded low scores in the word memorisation task as well, whereas the opposite was true for non-solfège subjects. The relationship, however, between Word span and Total task correct scores, was found non-significant for both conditions, in both groups.[2] In the taught solfège

---

[2] For the solfège group, $\tau = .09$, $p > .05$, *ns*, in the total sample and $\tau = -.09$, $p > .05$, *ns*, in the filtered sample over 70%; for the non-solfège group $\tau = .10$, $p > .05$, *ns*, in the total sample and $\tau = .39$, $p > .05$, *ns*, in the over 70% sample.

group there was only one participant out of seven who had a Total task correct score below 70; this subject had also produced a Word span score lower than the group mean, but the trend of low Total correct scores being accompanied by low Word span scores was not consistent within the taught group, either.

Regarding solfège groups' performance in the distraction task, the Kolmogorov-Smirnov test statistic was found non-significant for Total task correct scores, indicating that the distribution of scores was the same for the solfège and the non-solfège group. Mean Total task correct scores and their associated standard errors for all three groups, taught solfège, solfège and non-solfège, are presented in Table 5.15:

Total task correct (%) scores

	Mean	SE
Solfège	82.72	3.60
Non-solfège	84.86	3.03
Taught solfège	80.57	6.56

Table 5.15: Mean Total task correct scores for the three groups in Experiment B3.

The mean scores produced by subjects in the distraction task were very similar across the three groups (Table 5.16); moreover, a similar percentage of subjects in each group managed to score 70 or more in the task: 80% of participants in the solfège group, 84.20% from the non-solfège group and 6 out 7 (85.71%) from the taught solfège group managed to score 70 or above in the high–low task.

5.2.3.7. Summary

Hypothesis no. 1 stated that:

Solfège participants will be more likely to yield lower Word span scores compared to the non-solfège and the taught solfège group.

This hypothesis was not supported by data analysis; non-parametric tests as well as simple regression analysis yielded non-significant results regarding the effect of solfège group on Word span scores. It is interesting to note, however, that the effect of the solfège group on Word span performance was found borderline non-significant in the regression model (p = .058); this value, along with the reservations regarding the power of the distraction task to involve participants in mental engagement with note names, and given the relatively restricted size of the sample, raised certain concerns regarding the power of the experiment overall to detect interference effects of the pitch task on visual verbal recall. For these reasons, more thorough testing would be required in order to draw any definite conclusions regarding the exact way pitch labels function in memory.

Hypothesis no. 2 stated that:

Solfège participants will be more likely to yield lower Total task correct scores compared to non-solfège and taught solfège participants.

This hypothesis was not supported by data analysis, either: non-parametric tests as well as the observation of mean scores across the three groups indicated that the solfège, non-solfège and taught solfège groups had produced similar score patterns for the high–low task. No correlation was found in the sample between Word span and Total task correct scores.

5.2.4. Experiment B: Discussion

The aim of the three experiments in group B was to compare the effect of a pitch-comparison task on visual verbal memory in solfège and non-solfège musicians. For this purpose, a visual verbal memorisation task was run in parallel with a visual pitch comparison task; in experiment B1, the verbal memorisation task contained letters, in B2 it contained nonsense syllables and in B3 it contained monosyllabic words. The assumption behind the hypotheses posited in each of the three variants of Experiment B was that the pitch comparison task would employ pitch-naming strategies and therefore interfere with the performance of the memorisation task: non-solfège participants were expected to perform worse than solfège participants in Experiment B1, in which letters were used as memorisation stimuli, and solfège participants were expected to perform worse in Experiments B2 and B3, in which nonsense syllables and monosyllabic words, respectively, were used as memorisation stimuli. An important restriction in the design of the experiment was that the mode of performance of the distraction task was not imposed by the experiment itself, therefore it was not possible to ascertain that participants were actually using pitch naming strategies in order to perform the distraction task. Consequently, the interpretations elicited by data analysis had to be conservative, as it was not possible to be conclusive regarding the existence and the extent of the proposed effects.

The correlation of Span scores with Error scores for the experiments was found to be non-significant, apart from the case of the non-solfège group in Experiment B1; in all the other groups, performance in the memorisation task, whether memorising letters, syllables, or words, appeared to be statistically uncorrelated with performance in the distraction task. In the case of the non-solfège group in Experiment B1, Kendall's tau was found significant at $p > .05$, revealing a positive relationship between Letter span scores and Total task correct scores: non-solfège participants appeared to be more likely to perform better in the memorisation task if they did well in the distraction task. Despite the lack of correlation between scores in the majority of the cases, a decision was made to apply a 70% threshold on the performance of the distraction task: separate analyses were conducted for data from participants who had answered correctly in 70% or more of the high–low distraction task. The particular cut-off point was selected so that it would filter out participants who were performing close to chance in the

distraction task, while at the same time maintaining a sufficient sample size. It is important to note at this point that, given the correlation results across samples, participants who had not reached the cut-off point in the distraction task score should not be regarded as participants who did not complete the experiment in an appropriate manner; although this remained a possibility in some cases, the non-significance of correlation statistics indicated that the particular subjects should be better considered as subjects who were not as competent in completing the distraction task correctly, or were not able to do it accurately in parallel with the verbal task. This effect could be explained by a variety of reasons, such as lack of music-reading experience, lack of familiarity with the bass or treble clef or any other technical reason, rather than be directly attributed to lack of conformity with experiment rules.

Statistical analyses showed that solfège and non-solfège groups did not perform significantly differently in experiments B1 and B3, contrary to the expectations posited by experimental hypotheses. These findings suggested that, in Experiment B1, non-solfège subjects were not directly using letter names to perform the distraction task, so an interference effect with the memorisation task could not be traced; the same possibility is true for Experiment B3, where the performance of the distraction task without use of solfège note names by solfège participants could have led to the absence of any detrimental effect on memorisation and recall performance. In Experiment B2, however, the difference between the solfège and the non-solfège group was found to be statistically significant in the overall sample, but non-significant in the filtered sample of participants who had achieved over 70% in the distraction task. The participants eliminated in the filtered sample were 6 (13.64% of the total sample) and examination of their scores revealed that their group did not show internal consistency; three of the participants had yielded Syllable spans over their group mean and the remaining three had scored below the group mean. The characterisation, thus, of these participants as outliers presented problems; testing a larger sample would help clarify this ambiguity.

If, however, the decision to eliminate below-70% subjects was dismissed as arbitrary, what could be an explanation between the solfège and the non-solfège group results? Asserting the presence of interference between the memorisation and the distraction tasks was also problematic, because of the interpretation of results given in Experiments B1 and B3; it would be implausible to claim that solfège participants had used pitch

labels only for B2, but not for the other experiments. A possible solution could be provided by the definition of solfège note names as not strictly belonging to the realm of lexical items, but serving a distinct purpose in cognition (Chapter 1.2), which brought them closer to non-words. This would explain the lack of an interference effect in the solfège group for Experiment B3, in which the memorisation stimuli were words, but it would still be potentially problematic in explaining why non-solfège participants did not show the same effect in Experiment B1. A plausible explanation could yet again be derived from the theory of pitch labels developed in Chapter 1.2, in which it was argued that solfège note names, owing to their unique dual nature, are automatically used in every music-reading task; letter names, on the other hand, lacking the one-to-one relationship of signifier and signified with notes, are used by non-solfège individuals only when the need arises to refer to notes *in absentia*. This theory explained the absence of interference effects in Experiments B1 and B3, as well as the significant effect in B2, without the exclusion of any part of the tested sample; on the other hand, the problem of the non-significance of the over-70% sample analysis remained unresolved. A last attempt to clarify the ambiguity and shed more light on the possible effects of solfège use in memorisation tasks will be made in the last series of experiments, which will combine the unambiguous use of solfège with a verbal distraction task.

5.3. Experiment C

The aim of this series of experiments was to compare the effect of a verbal task on visual music memory in solfège and non-solfège musicians. In order to locate potential differences between solfège and non-solfège participants, the verbal task was designed with three levels:

4. **Phonemic level:** Subjects were presented with pairs of letters and asked to choose the letter which came first in alphabetical order.

5. **Morphemic level:** Subjects were presented with pairs of nonsense syllables and asked to choose the syllable whose consonant came first in alphabetical order. Despite this task being presented within a syllabic context, the task itself was letter-based, similar to the phonemic variation task; for this reason, it is perhaps more accurate to name this variation **morphemic/phonemic**, rather than purely morphemic.

6. **Lexical level:** Subjects were presented with pairs of monosyllabic words and asked to choose the word which contained the largest amount of consonants. Once more, the task, despite being presented within a word context, it was basically a letter-based, counting task; for this reason, a more accurate name for this variation would be to call it a **lexical/phonemic**.

In all three variants of the experiment, the verbal task was run in parallel with a music memorisation task.

5.3.1. Experiment C1: Phonemic interference in a music memorisation task

In Experiment C1, subjects were asked to memorise a 4-note sequence while simultaneously performing a letter task. The memorisation stimuli in Experiment C were identical to the stimuli used in Experiment A (section 5.1.1.3); the difference between the tests administered to solfège, taught solfège and non-solfège participants was the mode of recall: non-solfège subjects were provided with a list of letter note names, whereas solfège and taught solfège subjects were provided with a list of solfège note names. Taught solfège participants were expected to fall in the same two categories described in Chapter 5.1: a) participants performing a mental 'translation' between letter and solfège names during recall and b) participants using solfège names directly during the presentation of the memorisation stimulus. In both cases, taught solfège subjects were expected to have a relative disadvantage in performing the music memorisation task compared to solfège and non-solfège subjects. Moreover, the performance of the letter task stimulus was expected to interfere with the performance of the music memorisation task in the non-solfège group and in some subjects form the taught solfège group, who were using letter names to memorise the musical stimuli. In the taught solfège group, the interference of the letter distraction task and the assumed disadvantage resulting from the use of solfège note names were expected to have an accumulative negative effect on memorisation performance. According to these expectations, the following hypotheses were formulated:

1. **Non-solfège subjects will be more likely to yield lower Note span scores than solfège subjects.**

2. **Taught solfège subjects will be more likely to yield lower Note span scores than solfège and non-solfège subjects.**

3. **The distribution of correct answers in the distraction task will be similar across the three groups.**

As in experiments A and B (see Chapters 5.1 and 5.2), the taught solfège sample was extremely restricted in size, therefore statistical analyses were conducted using data from the solfège ad non-solfège groups only; comparisons with the taught solfège group were made by reference to score frequencies and mean scores.

5.3.1.1. Method: Design

The design of the experiment was a combination of the designs of the previous two experiments, A and B: it employed an independent design, using Solfège use as the independent variable, with three levels: solfège, taught solfège and non-solfège. The dependent variable was performance on the memorisation task, measured by the following parameters:

1. Note span, which was the total amount of notes remembered correctly.
2. Total task correct, which was the total amount of correct answers given in the distraction task.

5.3.1.2. Method: Participants

41 musicians in total completed the task: 22 were English and 19 were Greek. 4 of the English participants had received solfège training over the previous academic year and received the solfège version of the test; the remaining 18 English participants had never received formal training in solfège and were given the letter-name version of the test.

The English group included 14 men and 8 women, with a mean age of 25.41 years; 3 subjects were staff members and the rest were students. There was one AP possessor in the group; in terms of first-study instruments, there were 7 singers, 3 string players, 4 wind players, 6 pianists, 1 guitar player and 1 percussionist.

The Greek group included 2 men and 17 women, with a mean age of 21.21 years; all participants in this experiment were students. 2 subjects were AP possessors and 1 had reported she did not know whether or not she had AP; the remaining participants were

not AP possessors. There were 2 string players, 1 wind player, 13 pianists and 1 guitar player in the Greek group.

5.3.1.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were the same as the stimuli for Experiment A (section 5.1.1.2). Each stimulus was presented in the form of a 4/4 bar containing a 4-crotchet sequence, using only natural notes, with no repeats. Each melody was presented only visually and lasted 5 seconds on the screen. For the recall of the sequence, participants were presented with a screen containing a list of all notes with their letter or solfège names, according to whether they belonged in the solfège, the taught solfège, or the non-solfège group; from this list of options, they were asked to select the correct notes in the correct order they had occurred. In addition to the seven note names, a set of four 'Blank' options was available, in case a participant forgot or was dubious about the pitch of a note. There was no time restriction in the recall screen. After participants made their selection and clicked on 'Next', a feedback screen for the memorisation task appeared for 2 seconds, informing participants of the number of notes they had remembered correctly.

5.3.1.4. Materials and Procedure: Letter task stimuli

Between the presentation of the memorisation stimulus and its recall, subjects were asked to perform a verbal comparison task. For this task, they were presented with pairs of letters (e.g. T, Q; O, V) ; letters of each pair appeared in sequence and stayed on screen for 2 seconds. During the 2 seconds, subjects had to choose if the 1st or the 2nd letter, in order of presentation, came first in the alphabet, by pressing 1 or 2, respectively, on the numeric keypad. If a subject failed to press any key during the 2 seconds, the experiment moved on to the next letter pair, counting the non-response as an incorrect answer. A feedback screen appeared for 1 second after each letter task, informing participants whether or not they had got the answer right.

In the English version of the experiment, letter pairs in the verbal comparison task did not contain any of the first 8 letters in the English alphabet (A, B, C, D, E, F, G, H), in order to avoid confusion with note letter names. In the Greek version of the experiment, the letters A, B, E and H of the Greek alphabet were not used, for the same reason. The importance of performing the letter task as quickly and as accurately as possible was repeatedly stressed, both in the written and in the verbal instructions provided.

5.3.1.5. General procedure

As in all previous experiments on E-Prime, participants read extensive instructions and performed practice trials in order to get familiarised with the procedure, before proceeding to the main experiment. The practice trials for Experiment C followed the same structure as in Experiments A and B: at first, subjects performed the music memorisation task separately, then they performed the verbal task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different note sequence was presented, which the subject was asked to memorise; between the presentation of the memorisation stimulus and its recall, the verbal task took place. The number of distraction tasks performed in each experimental set varied between 2–7; the number of distraction tasks performed during the three experimental sets varied between 11–15. A sample procedure of a single experimental set for the non-solfège group is shown in Figure 5.12:

Figure 5.12: A sketch of the procedure followed in each of the three experimental sets in Experiment C1. Note that stages 2, 3 and 4 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Note span, which is equal to the total number of notes remembered correctly in the correct position, across all 3 experimental sets. The maximum score was 12, which is the sum of all notes in the three 4-note sequences used in the memorisation tasks.

- Total errors, which is equal to the total number of errors made in the parallel verbal task. This score is further broken down in two types of errors:

  - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

  - Accuracy errors, which is the total number of errors resulting from the subject answering incorrectly to the letter task.

In addition to these scores, the number of correct answers given in the distraction task was also used in the analysis; it was calculated by subtracting the Total errors score from the total number of letter tasks performed throughout the 3 experimental sets. This was the **Total task correct** score and was calculated only as a percentage score.

5.3.1.6. Results

Table 5.16 displays the descriptive statistics for the memorisation and the distraction task scores separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	Median	SD	Min.	Max.	Skew	Kurtosis
Note span	Solfège	11.11	12	1.49	8	12	- 1.33	0.16
	Taught solfège	11.25	11.50	.96	10	12	- 0.86	- 1.29
	Non-solfège	10.44	12	3.01	0	12	- 2.88	8.93
Note span (%)	Solfège	92.54	100	12.39	66.67	100	- 1.33	0.16
	Taught solfège	93.75	95.83	7.98	83.33	100	- 0.86	- 1.29
	Non-solfège	87.04	100	25.12	0	100	- 2.88	8.93
Total letter errors	Solfège	2.79	3	1.81	0	6	0.41	- 0.86
	Taught solfège	3.75	4	2.22	1	6	- 0.48	- 1.70
	Non-solfège	4	3	3.46	0	15	2.01	5.34
Speed errors	Solfège	0.37	0	0.50	0	1	0.59	- 1.86
	Taught solfège	0.25	0	0.50	0	1	2	4
	Non-solfège	1.06	0	3.51	0	15	4.14	17.38
Accuracy errors	Solfège	2.42	2	1.84	0	6	0.56	- 0.37
	Taught solfège	3.5	3.5	2.38	1	6	0	- 4.34
	Non-solfège	2.94	2	2.26	0	7	0.52	- 0.70
Total task correct (%)	Solfège	79.71	80	12.97	58.33	100	- 0.37	- 1.08
	Taught solfège	72.92	70	16.80	58.33	93.33	0.51	- 3.06
	Non-solfège	72.15	80	22.99	0	100	- 1.90	4.98

Table 5.16: Descriptive statistics for C1 test scores, by solfège knowledge group. It is important to note the difference in group sizes: for the solfège and the non-solfège group N = 19 and N = 16, respectively, whereas for the taught solfège group N = 4.

Values which could cause concern in Table 5.17 could be found in the Note span and Total errors scores for the non-solfège groups: for Note span the minimum value was zero, meaning that at least one subject had recalled no notes correctly, in any of the sequences, and in Total errors and Total speed errors the maximum value was 15, which was also the maximum number of distraction tasks in the experiment, implying that there was at least one subject who had effectively failed to complete the distraction task in parallel with the memorisation. Box plots were examined in order to find out whether these scores belonged to the same subject, as this would mean that this subject had failed to complete the experiment; the outlier scores, however, were found to belong to three different participants, all coming from the non-solfège group: one participant had failed to recall correctly any notes in the memorisation task and two participants had produced zero answers in the letter task. Although data from the zero Note span participant were excluded in some of the following analyses, interpreted as belonging to an exceptionally low-span participant, zero Note span was still considered to be an acceptable score within the current experimental design. On the other hand, zero correct answers in the distraction task, especially considering that all errors were speed errors, was considered to be a score indicating that the participant had failed to complete the distraction task altogether. As the purpose of the experiment was to test performance on the memorisation task monitoring the interference caused by the parallel task, data from the zero Total task correct participants were excluded from all analyses; for all subsequent analyses, the number of subjects in the non-solfège group is N = 16, unless indicated otherwise.

Note span and Total task correct scores were plotted, both collectively and separately for the solfège and non-solfège groups. All Note span scores appeared to have significantly non-normal distributions, whereas this was not the case with distraction task scores, as shown in Table 5.17:

Tests of Normality

	Note span			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.643	.667	.693	.929	.910	.925
df	19	16	39	19	16	39
Sig.	.000	.000	.000	.169	.118	.012

Table 5.17: Values for the Shapiro-Wilk test statistic are shown for the distribution of scores in the two groups separately as well as in the whole sample, which included solfège, taught solfège and non-solfège participants. All Note span distributions were found significantly non-normal at $p < .001$; subgroup distributions of Total task correct scores were found normal, but the combined scores appeared to create a significantly non-normal distribution ($p < .05$).

Because of the violation of the normality assumption, non-parametric tests were conducted on Note span data. The Kolmogorov-Smirnov Z test statistic was found non-significant for Note span in the solfège and the non-solfège groups, indicating that the distribution of Note span scores was the same between the two groups; the Moses extreme reactions test, however, was found significant at $p < .01$, indicating that the variability of scores was different in each group. The same results were found for Note span scores in the filtered sample, including only participants who had scored 70 or above in the distraction task: the distribution of scores appeared similar for the two groups, whereas the Moses extreme reactions test was found significant at $p < .05$, indicating there was a significant difference in the range of scores in each group. A graph displaying score frequencies for each group in the over 70% sample is shown in Figure 5.13:
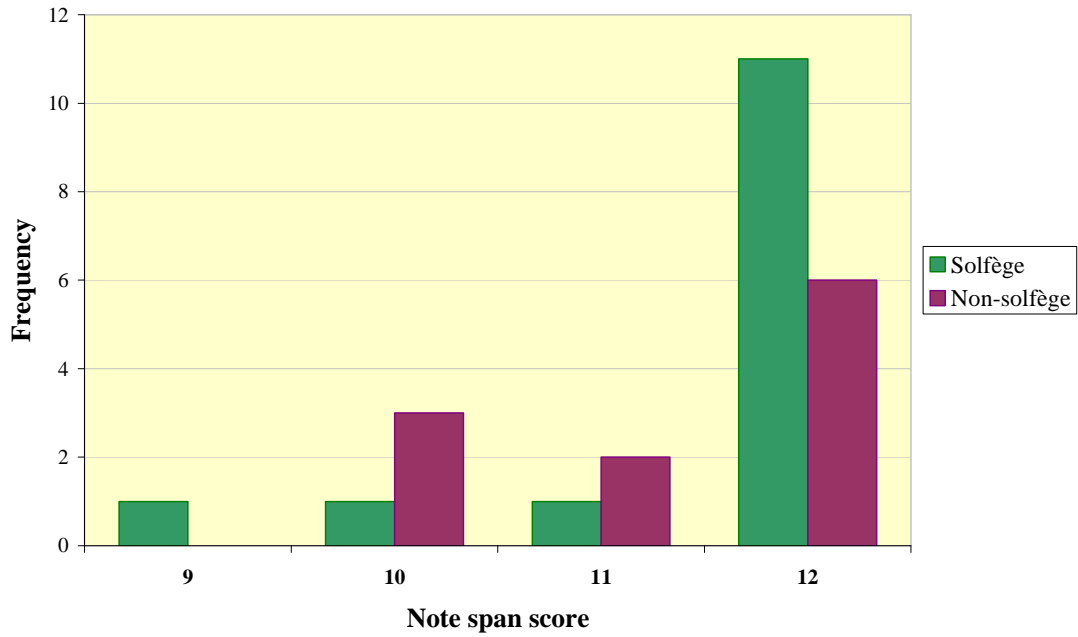
Figure 5.13: Note span frequencies for the solfège and the non-solfège groups, including only participants who had scored 70% correctly or above in the distraction task. The maximum possible score was 12; group sizes were N = 14 for the solfège and N = 11 for the non-solfège group.

Figure 5.13 shows that 11 out of 14 solfège participants who scored over 70% correctly in the distraction task recalled all 12 notes in the sequences correctly, whereas the equivalent figure for the non-solfège group was 6 out of 11 participants. The graph is useful in understanding the significance values appearing in the non-parametric tests described above: the similarity of the scores' distribution can be attributed to the negative skewness in both distributions, as well as to similar means and standard deviations for both groups;[3] a similar pattern of results could be observed in the graph of the overall sample, which had also yielded a non-significant Kolomogorov-Smirnov and a significant Moses extreme reactions test statistic. In the taught solfège group, there were 2 subjects who had scored over 70% in the distraction task, who had also

---

[3] Descriptive statistics for Note span score in the over 70% Total task correct score sample were the following:

	Mean	Median	SD	Skewness	Kurtosis
Solfège	11.57	12	0.94	- 2.20	4.12
Non-solfège	11.27	12	0.91	- 1.55	- 1.55

both scored 100% (12 notes) in the memorisation task; the other two subjects, who had scored below 70% in the letter task, had a Note span of 10 and 11, respectively.

Regarding participants' performance in the letter task, no relationship was found between solfège knowledge and the Total task correct score, $r_b = -0.01$, $p > .05$, suggesting that solfège and non-solfège participants had yielded a similar pattern of performances in the distraction task. Mean scores for the three groups and their associated standard errors are displayed in Table 5.18:

Total task correct (%) scores

	Mean	**SE**
Solfège	79.71	2.98
Non-solfège	76.48	5.74
Taught solfège	72.92	8.40

Table 5.18: Mean Total task correct scores for the solfège, non-solfège and taught solfège groups.

In total, 5 participants (26.32%) from the solfège and another 5 (31.25) from the non-solfège group failed to reach the 70% threshold in the letter task. No significant correlation was found between the number of errors in the letter task and the amount of notes remembered correctly in the correct position.

5.3.1.7. Summary

Hypothesis no. 1 stated that:

Non-solfège subjects will be more likely to yield lower Note span scores than solfège subjects.

This statement was not supported by findings from the experiment C1; solfège and non-solfège group means for Note span were very similar (11.57 and 11.27 notes,

respectively, for participants who had scored over 70 in the distraction task) and non-parametric tests conducted on the data did not reveal a significant difference between the groups.

Hypothesis 2 stated that:

Taught solfège subjects will be more likely to yield lower Note span scores than solfège and non-solfège subjects.

For this experiment, the taught solfège sample included only 4 participants, so it was impossible to draw any conclusions based on comparisons from the particular dataset; it is also worth noting that two out of the four subjects failed to reach the 70% threshold in the distraction task, rendering the comparison with the other 2 groups even harder. Despite the limitations of the sample, it should be mentioned that there was no evidence of support for the hypothesis, as the Note span mean for the taught solfège group was slightly higher than that of the other two groups.

The third hypothesis stated that:

The distribution of correct answers in the distraction task will be similar across the three groups.

This hypothesis was supported by data analysis; non-parametric test statistics suggested that the distribution of correct answers in the letter task was similar for the solfège and the non-solfège groups, while the mean Total task correct score in the taught solfège group was very close to the other two groups' means.

5.3.2. Experiment C2: Morphemic/phonemic interference in a music memorisation task

Experiment C2 was identical to Experiment C1, apart from the stimuli in the distraction task: instead of being presented with single letters, subjects were presented with nonsense syllables. As with Experiments B2 and B3, nonsense syllables in C2 were used in order to compare their effect with that of monosyllabic words in C3; differences in the extent of interference caused by the syllable task on the recall task was expected to illuminate the nature, function and use of solfège syllables. In C2, the reading of syllables required to perform the distraction task was expected to interfere with the recall of the pitch sequences in the solfège and the taught solfège group; consequently, the following hypotheses were tested:

1. **Note span scores will be lower in the solfège and the taught solfège groups compared to the non-solfège group.**

2. **The distribution of correct answers in the distraction task will be similar across the three groups.**

5.3.2.1. Method: Design

The design of Experiment C2 was identical to that of Experiment C1, described in section 5.3.1.1.

5.3.2.2. Method: Participants

A total of 41 subjects took part in this experiment; 20 subjects were tested in Greece and 21 subjects were tested in England. Out of the English participants, 4 belonged in the taught solfège group and completed the solfège version of the English experiment, while the other 17 English participants completed the letter-name version of the experiment.

There were 14 men and 7 women in the English group, with a mean age of 21.8 years; 2 participants from the group were music staff members and the rest were music students. There was one subject in the group who had AP and one subject who did not know whether he had AP; all the remaining participants did not have AP. The English group consisted of 4 singers, 5 string players, 2 wind players, 6 pianists, 3 guitar players and 1 percussionist.

The Greek group included 5 men and 15 women, with a mean age of 22 years; there was one staff member amongst participants and the rest were all music students. According to questionnaire responses there were no AP possessors in the Greek group for this experiment; there was, however, a participant who had not responded to the AP question. Instrumental groups included 2 singers, 2 string and 2 wind players, 12 pianists and 2 guitar players.

5.3.2.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were the same as stimuli for Experiment A: each stimulus was presented visually, for a duration of 5 seconds, in the form of a 4/4 bar containing a 4-crotchet sequence. For the recall of the sequence, participants were presented with a screen containing a list of all notes with their letter or solfège names; from this list of options, they were asked to select the correct notes in the correct order they had occurred. In addition to the seven note names, a set of four 'Blank' options was available. No time restriction was imposed on the recall screen; after participants had made their selection and clicked on 'Next', a feedback screen for the memorisation task appeared for 2 seconds, informing participants of the number of notes they had remembered correctly.

5.3.2.4. Materials and Procedure: Syllable task stimuli

Between the presentation of the memorisation stimulus and its recall, subjects were asked to perform a verbal comparison task. For this task, they were presented with pairs

of nonsense syllables (e.g. ke, af, uw); all syllables used in the task fulfilled the following criteria:

1. They consisted of 1 consonant and 1 vowel, in either order.
2. They did not form abbreviations or acronyms in themselves.
3. The serial combination of syllables in each pair did not form words, abbreviations or acronyms.
4. The syllables in each pair could include the same vowel, but could not include the same consonant.

Syllables in each pair appeared in sequence and stayed on the screen for 2 seconds; during this time, subjects had to choose if the consonant of the 1st or the 2nd syllable, in order of presentation, came first in the alphabet, by pressing 1 or 2, respectively. If a subject failed to press any key during the 2 seconds, the experiment moved on to the next syllable pair, counting the non-response as an incorrect answer. A feedback screen appeared for 1 second after each syllable pair, informing participants whether or not they had got the answer right.

5.3.2.5. General procedure

As in all previous experiments on E-Prime, participants were provided with explicit instructions and performed extensive practice trials in order to get familiarised with the procedure, before proceeding to the main experiment. The practice trials followed the same structure as in the previous experiments: at first, subjects performed the music memorisation task separately, then they performed the verbal task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different note sequence was presented, which the subject was asked to memorise; between the presentation of the memorisation stimulus and its recall, the verbal task took place. The number of distraction tasks performed in each experimental set varied between 2–7; the number of distraction tasks performed during the three experimental sets varied between

11–15. A sample procedure of a single experimental set for the non-solfège group is shown in Figure 5.14:

Figure 5.14: A sketch of the procedure followed in each of the three experimental sets in Experiment C2. Note that stages 2, 3 and 4 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Note span, which is equal to the total number of notes remembered correctly in the correct position, across all 3 experimental sets. The maximum score was 12, which is the sum of all notes in the three 4-note sequences used in the memorisation tasks.

- Total errors, which is equal to the total number of errors made in the parallel verbal task. This score is further broken down in two types of errors:
    - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

- ▪ Accuracy errors, which is the total number of errors resulting from the subject answering incorrectly to the syllable task.

In addition to these scores, the number of correct answers given in the distraction task was also used in the analysis, called **Total task correct** and calculated only as a percentage score.

5.3.2.6. Results

Descriptive statistics for scores in the memorisation and the distraction tasks are presented in Table 5.19 separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	Median	SD	Min.	Max.	Skew	Kurtosis
Note span	Solfège	10.65	11.50	1.84	6	12	- 1.27	0.57
	Taught solfège	10.25	10	1.26	9	12	1.13	2.23
	Non-solfège	10.82	11	1.38	7	12	- 1.42	2.37
Note span (%)	Solfège	88.75	95.83	15.36	50	100	- 1.27	0.57
	Taught solfège	85.41	83.33	10.49	75	100	1.13	2.23
	Non-solfège	90.20	91.67	11.50	58.33	100	- 1.42	2.37
Total syllable errors	Solfège	3.05	3	2.09	0	7	0.23	- 0.98
	Taught solfège	3.50	3.50	1.29	2	5	0	- 1.20
	Non-solfège	4.71	4	3.58	1	15	1.84	4.01
Speed errors	Solfège	0.20	0	0.41	0	1	1.62	0.70
	Taught solfège	0.25	0	0.50	0	1	2	4
	Non-solfège	1	0	3.62	0	15	4.07	16.66
Accuracy errors	Solfège	2.85	3	1.95	0	7	0.23	- 0.76
	Taught solfège	3.25	3.50	1.71	1	5	- 0.75	0.34
	Non-solfège	3.71	4	2.33	0	9	0.67	0.41
Total task correct (%)	Solfège	77.76	78.46	15.13	41.67	100	- 0.51	0.07
	Taught solfège	72.87	75.13	13.45	54.55	86.67	- 0.95	1.70
	Non-solfège	66.73	73.33	22.94	0	93.33	- 1.68	3.70

Table 5.19: Descriptive statistics for C2 test scores for the three groups. Group sizes were N = 20 and N = 17 for the solfège and the non solfège group, respectively, and N = 4 for the taught solfège group.

As in Experiment C1 (Table 5.16), there appeared to be a subject in the non-solfège group who failed to produce any response in the whole of the distraction task; this subject's data were excluded from all subsequent analyses.[4] After eliminating the outlier, the Shapiro-Wilk test was conducted on the Note span and Total task correct scores, both collectively and separately for the solfège and non-solfège groups. As shown in Table 5.20, Note span scores appeared significantly non-normal, whereas the distribution of Total task correct scores appeared normal for all groups:

Tests of Normality

	Note span			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.756	.807	.807	.960	.951	.970
df	20	16	40	20	16	40
Sig.	.000	.000	.000	.541	.512	.352

Table 5.20: Values for the Shapiro-Wilk test statistic. Distributions of Note span scores appear significantly non-normal for both subgroups as well as for the overall sample, whereas distraction task scores appear normal.

Because of the violation of the normality assumption, non-parametric tests were selected for the analysis of Note span scores. The Kolmogorov-Smirnov Z test statistic was found non-significant for Note span in the solfège and the non-solfège groups, indicating that the distribution of Note span scores was the same between the two groups; the Moses extreme reactions test, however, was found significant at p < .05, indicating that the variability of scores was different in each group. The Moses test was not significant in the filtered, over-70% correct in the distraction task sample; it is important to note that a total of 14 subjects from the solfège group and 9 subjects from the non-solfège group scored over the 70% threshold. Figure 5.15 displays the frequencies for Note span scores in the unfiltered and the unfiltered sample, separated by solfège group:

---

[4] Note that it was a different participant in Experiments C1 and C2 who failed to perform the distraction task.
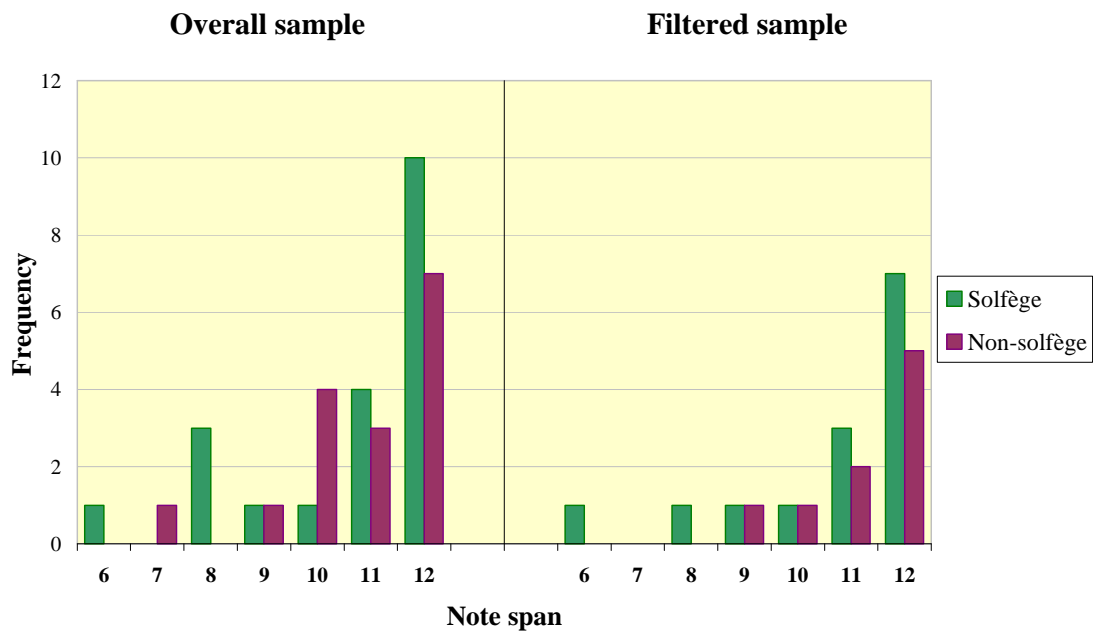
Figure 5.15: Note span frequencies for the solfège and the non-solfège group, in the overall sample, including all solfège and non-solfège subjects and in the filtered sample, including only subjects who scored 70% or above in the distraction task. The maximum possible score was 12 notes; group sizes were N = 36 for the overall sample and N = 23 for the over 70% sample.

Figure 5.15 shows that the pattern of performances was very similar between the filtered and the unfiltered sample, indicating that the different Moses extreme reactions test were a result of the small sample sizes, combined with the large number of tied scores. Out of the 4 taught solfège participants, one had failed to reach the 70% threshold in the distraction task and had yielded a Note span of 10; the remaining three participants had all scored above 70% correctly in the syllable task and had yielded scores equal to 9, 10 and 12 notes, respectively.

Regarding participants' performance in the syllable task, no correlation was found between Total task correct scores and the solfège group the participants belonged in, r_b = 0.28, p > .05; mean scores and associated standard errors for the three solfège groups are displayed in Table 5.21:

Total task correct (%) scores

	Mean	SE
Solfège	77.76	3.38
Non-solfège	70.90	5.56
Taught solfège	72.87	6.73

Table 5.21: Mean Total task correct scores for the solfège, non-solfège and taught solfège groups.

In total, 6 participants (30%) from the solfège group had failed to reach the 70% threshold in the letter task, whereas the equivalent number for the non-solfège group was 7 participants (43.8%). No significant correlation was found between the number of errors in the distraction task and the amount of notes remembered correctly in the correct position.

5.3.2.7. Summary

Hypothesis no. 1 stated that:

Note span scores will be lower in the solfège and the taught solfège groups compared to the non-solfège group.

This hypothesis was not supported by findings; non-parametric tests showed that the distribution of note span scores was the same across the solfège and non-solfège groups, while the mean Note span in the taught solfège group was very close to that of the other two groups; the exclusion of participants who had scored below 70% correctly in the distraction task did not alter these results.

The second hypothesis stated that:

The distribution of correct answers in the distraction task will be similar across the three groups.

This hypothesis was supported by data analysis; after the removal of an outlier, who had produced zero answers in the syllable task, the distribution of correct answers was the same in the solfège and non-solfège groups. It is worth noting, however, that the percentage of participants failing to reach the 70% threshold was particularly high in Experiment C2 (30% for the solfège and 43.8% for the non-solfège group); performance in the distraction task across experiments will be discussed in Chapter 5.4.

5.3.3. Experiment C3: Lexical/phonemic interference in a music memorisation task

Experiment C3 was identical to Experiments C1 and C2, apart from the stimuli in the distraction task: instead of being presented with letters or nonsense syllables, subjects were presented with lexical phonemes, i.e. monosyllabic words. As in Experiment C2, the performance of the word-based distraction task was expected to affect negatively performance in the memorisation task for subjects who used solfège note names. Consequently, the following hypotheses were tested:

1. **Solfège and taught solfège participants will have lower Note span scores compared to non-solfège participants.**

2. **The distribution of correct answers in the distraction task will be similar across the three groups.**

5.3.3.1. Method: Design

The design of Experiment C was identical to C1 and C2: Solfège use was the independent variable, with three levels: solfège, taught solfège and non-solfège, and performance on the memorisation task was the dependent variable, measured by the following parameters:

3. Note span, which was the total amount of notes remembered correctly.
4. Total task correct, which was the total amount of correct answers given in the distraction task.

5.3.3.2. Method: Participants

A total of 41 subjects took part in this experiment; 19 subjects were tested in Greece and 22 subjects were tested in England. Out of the English participants, 5 belonged in the taught solfège group and completed the solfège version of the English experiment,

while the other 17 English participants completed the letter-name version of the experiment.

13 men and 9 women participated in the English group, with a mean age of 25.36 years; 2 participants from the group were music staff members and the rest were music students. There was one subject in the group who had AP and 3 subjects who did not know whether or not they had AP; all the remaining participants were not AP possessors. The English group consisted of 3 singers, 5 string players, 4 wind players, 7 pianists and 3 guitar players.

The Greek group included 6 men and 13 women, with a mean age of 21.05 years; all subjects in this group were students. According to questionnaire responses there were 2 AP possessors in the group and one person had failed to respond to the AP question; the rest were non-AP subjects. Instrumental groups included 4 singers, 2 string and 2 wind players, 9 pianists and 2 guitar players.

5.3.3.3. Materials and Procedure: Memorisation task stimuli

The memorisation task stimuli were the same as in the previous experiments in group C: each stimulus was a 4-crotchet sequence, presented visually for 5 seconds. For the recall of the sequence, participants were presented with a screen containing a list of all notes with their letter or solfège names; from this list of options, they were asked to select the correct notes in the correct order they had occurred. In addition to the seven note names, four 'Blank' options were also provided. No time restriction was imposed on the recall screen; after participants had made their selection and clicked on 'Next', a feedback screen for the memorisation task appeared for 2 seconds, informing participants of the number of notes they had remembered correctly.

5.3.3.4. Materials and Procedure: Verbal task stimuli

Between the presentation of the memorisation stimulus and its recall, subjects were asked to perform a verbal comparison task. For this task, they were presented with pairs of monosyllabic words: words in each pair appeared in sequence and stayed on the screen for 2 seconds. During the presentation of the second word, subjects had to choose whether the 1st or the 2nd word, in order of presentation, contained the biggest amount of consonants. If a subject failed to press any key during the 2 seconds, the experiment moved on to the next word pair, counting the non-response as an incorrect answer. A feedback screen appeared for 1 second after each word pair, informing participants whether or not they had got the answer right.

5.3.3.5. General procedure

As in all previous experiments on E-Prime, participants were provided with explicit instructions and performed extensive practice trials in order to get familiarised with the procedure, before proceeding to the main experiment. The practice trials followed the same structure as in the previous experiments: first, subjects performed the music memorisation task separately, then they performed the verbal task separately and, at the end, they performed two practice trials combining the two parallel tasks, in an exact simulation of the experimental procedure.

The core procedure comprised three experimental sets. In each set, a different note sequence was presented, which the subject was asked to memorise; between the presentation of the memorisation stimulus and its recall, the verbal task took place. The number of distraction tasks performed in each experimental set was between 2–7; the number of distraction tasks performed during the three experimental sets was between 11–15. A sample procedure of a single experimental set for the non-solfège group is shown in Figure 5.16:

Figure 5.16: A sketch of the procedure followed in each of the three experimental sets in Experiment C3. Note that stages 2, 3 and 4 were repeated between 2 and 7 times.

At the end of the experiment, the program reported the following scores, both in absolute and in percentage values:

- Note span, which is equal to the total number of notes remembered correctly in the correct position, across all 3 experimental sets. The maximum score was 12, which is the sum of all notes in the three 4-note sequences used in the memorisation tasks.

- Total errors, which is equal to the total number of errors made in the parallel verbal task. This score is further broken down in two types of errors:

    - Speed errors, which is the total number of errors reported by the experiment due to failure of the subject to press any key within the 2-second timeframe.

5.3.3.6. Results

Descriptive statistics for scores in the memorisation and the word tasks are presented in
Table 5.22 separately for the three groups: solfège, taught solfège and non-solfège:

		Mean	Median	SD	Min.	Max.	Skew	Kurtosis
Note span	Solfège	11.68	12	0.48	11	12	- 0.86	- 1.42
	Taught solfège	9.6	9	2.30	7	12	0.20	- 2.18
	Non-solfège	10.12	11	2.40	3	12	- 1.86	3.99
Note span (%)	Solfège	97.37	100	3.98	91.67	100	- 0.86	- 1.42
	Taught solfège	80	75	19.18	58.33	100	0.20	- 2.18
	Non-solfège	84.31	91.67	19.96	25	100	- 1.86	3.99
Total word errors	Solfège	2.53	1	3.63	0	12	2.10	3.78
	Taught solfège	2	3	1.41	0	3	- 0.88	- 1.75
	Non-solfège	1.94	0	1.64	0	6	0.98	0.85
Speed errors	Solfège	0.79	0	2.74	0	12	4.23	18.18
	Taught solfège	0	0	0	0	0		
	Non-solfège	0.18	0	0.39	0	1	1.87	1.67
Accuracy errors	Solfège	1.74	1	2.88	0	12	2.81	9.12
	Taught solfège	2	3	1.41	0	3	- 0.88	- 1.75
	Non-solfège	1.76	1	1.52	0	5	0.81	- 0.34
Total task correct (%)	Solfège	81.53	93.33	27.46	0	100	- 2.22	4.54
	Taught solfège	86.46	80	9.26	80	100	0.98	- 1.19
	Non-solfège	86.37	86.67	11.28	60	100	- 0.81	0.22

Table 5.22: Descriptive statistics for C3 test scores for the three groups. Group sizes were N = 19 and N
= 17 for the solfège and the non solfège group, respectively, and N = 4 for the taught solfège group.

Table 5.22 showed that there was at least one participant in the solfège group who had failed to produce any answers in the distraction word task, yielding a zero Total task correct score.[5] Data were examined and showed that there was one participant in the solfège group who had not completed the distraction task; data from this subject were removed from all further analyses. Following the removal of the outlier, the Shapiro-Wilk test was conducted on Note span and Total task correct scores for all groups; the tests revealed that all distributions except one were significantly non-normal. Table 5.23 shows normality test statistics for the solfège and the non-solfège subgroups, as well as for the total sample:

Tests of Normality

	Note span			Total task correct (%)		
	Solfège	**Non-solfège**	**All**	**Solfège**	**Non-solfège**	**All**
Shapiro-Wilk	.601	.782	.691	.690	.915	.770
df	18	17	40	18	17	40
Sig.	.000	.001	.000	.000	.122	.000

Table 5.23: Values for the Shapiro-Wilk test statistic. All distributions were found non-normal, apart from the Total task correct score for the non-solfège group.

Because of the violation of the normality assumption, non-parametric tests were selected for the analysis of Note span scores. The Mann-Whitney U as well as the Moses extreme reactions tests were both found significant, at $p < .05$ and $p < 001$, respectively, indicating that the distribution of scores across groups, as well as the variability of scores within each group, were different for the solfège and the non-solfège group. The Kolmogorov-Smirnov Z test statistic, on the other hand, was found non-significant at $p > .05$; although the latter is considered to have better power in

---

[5] The corresponding speed errors for this subject were 12 because, as explained in section 5.2.1.5, the number of distraction tasks across the three experimental sets could vary between 11 and 15; the subject who failed to complete the distraction task in this instance was a subject who had received a test having 12 distraction tasks overall.

smaller samples, the discrepancy between the two tests and the significance of the Moses test justified further examination of the data.

Simple regression analysis was conducted on Note span scores, using solfège knowledge as the predictor variable with two levels: solfège and non-solfège; as the assumption of normally distributed errors had been violated, bootstrap samples were used for the analysis. Regression results are displayed in Table 5.24:

Note span scores

	B	**SE B**[a]	**β**
Constant	10.12	.41	
Solfège knowledge	1.55	.58	.42**

$R^2 = .18$, p < .05

** p < .05, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap samples.

Table 5.24: Simple regression results for Note span scores.

The value of R^2 indicates that solfège knowledge affected the model significantly as a predictor; the value of β, however, suggested that, contrary to predictions, solfège subjects had yielded significantly higher Note span scores than non-solfège subjects. The same regression analysis was conducted on data from participants who had scored over 70% correctly in the word distraction task; 3 participants had been eliminated from the dataset as a result of the filtering procedure. Regression results for the over-70% sample were also found significant at p < .05,[6] validating further the conclusion that

---

[6] Regression results for the over 70% correct sample were as follows:

	B	**SE Ba**	**β**
Constant	10.31	0.54	
Solfège knowledge	1.44	0.55	.41**

$R^2 = .16$, p < .05; ** p < .05, *s*

a. Bootstrapped values. Significance for β is also based on bootstrapped samples.

solfège participants were significantly more likely to produce higher Note span scores than their non-solfège counterparts. In the taught solfège group, all five subjects had got over 70% of the answers in the distraction task correctly; their mean Note span score was lower than mean Note span scores for the solfège and the non-solfège group, in the filtered sample (Figure 5.17):

Mean Note span score
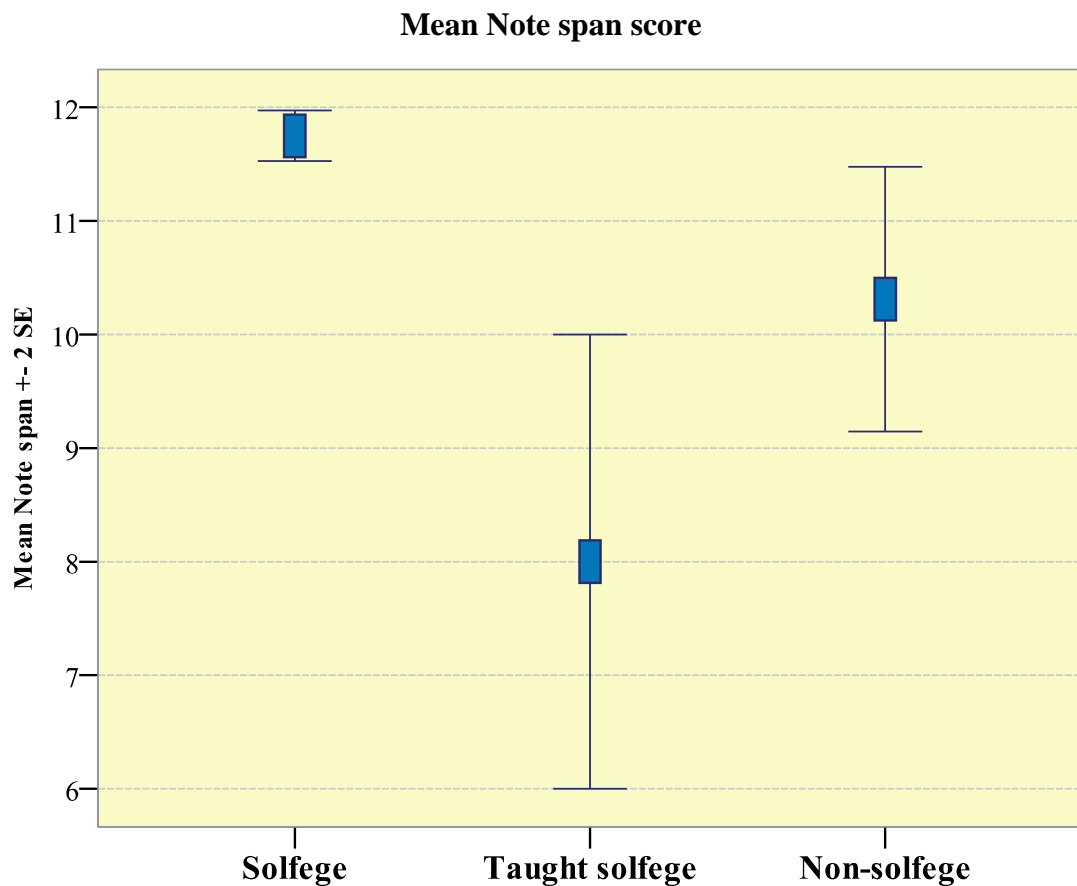
Figure 5.17: Mean Note span scores of the three solfège groups, for participants who had scored over 70% correctly in the distraction task. For the Solfège group (N =) M = 11.75, for the taught solfège group (N =) M = 9.6 and for the non-solfège group (N =) M = 10.31.

Regarding participants' performance in the letter task, all non-parametric test statistics were found non-significant, suggesting that solfège and non-solfège participants had yielded a similar pattern of performances in the distraction task. Mean scores for the three groups with their associated standard errors are displayed in Table 5.25:

Total task correct (%) scores

	Mean	SE
Solfège	86.06	4.63
Non-solfège	86.37	2.74
Taught solfège	86.46	4.14

Table 5.25: Mean Total task correct scores for the solfège, non-solfège and taught solfège groups.

In total, 2 participants (11.11%) from the solfège group had failed to reach the 70% threshold in the letter task, whereas the equivalent number for the non-solfège group was 1 participant (5.88%); all taught solfège participants had scored above 70 in the distraction task. No significant correlation was found between the number of errors in the word task and the amount of notes remembered correctly in the correct position in the overall sample; within solfège subgroups, however, a significant positive correlation was found between the two scores only in the solfège group, $\tau = .61$, $p < .01$.

5.3.3.7. Summary

Hypothesis no. 1 stated that:

Solfège and taught solfège participants will have lower Note span scores compared to non-solfège participants.

This hypothesis was not supported by data analysis; on the contrary, simple regression analysis results showed that solfège participants were more likely to yield higher Note span scores than non-solfège participants. For the solfège group only, a significant correlation was also found between distraction task performance and memorisation performance. On the other hand, taught solfège subjects had the lowest mean Note span score in the three groups; this finding, however, could not be generalised as it was derived from data from only 5 participants.

The second hypothesis stated that:

> The distribution of correct answers in the distraction task will be similar across the three groups.

This hypothesis was supported by analysis results; following the removal of the outlier, who had produced zero responses in the distraction task, the distribution of correct answers was the same in the solfège and non-solfège groups. It is also worth noting that the percentage of participants scoring above 70% in the distraction task in C3 was the highest compared to C1 and C2 results; participant performance in the distraction task will be discussed in the following chapter.

5.3.4. Experiment C: Discussion

The aim of the experiments in group C was to compare the effects of various verbal tasks on visual music memory, in solfège and non-solfège musicians. For this purpose, subjects had to memorise a short note sequence whilst performing a verbal task in parallel: in Experiment C1 the verbal task was letter-based, in Experiment C2 it was syllable-based and in Experiment C3 there was a word-based distraction task. The assumption behind the hypotheses posited in each of the three variants of Experiment C was that, as the recall screen prompted participants to use note names in order to identify the memorised sequence, the names would interfere selectively with the verbal task: non-solfège participants, using letter names to memorise the notes, were expected to perform worse than their solfège counterparts in C1, due to the parallel performance of the letter-based task; recall in solfège participants, on the other hand, was expected to be negatively affected by the parallel performance of syllable- and word-based tasks in C2 and C3, respectively. Taught solfège participants were expected to experience difficulty in all cases, as they were assumed to be either performing a 'translation' between letter and solfège names or attempting to memorise the note sequence using the less familiar solfège names. A restriction in the design of the experiment was that, although in C2 and C3 subjects were presented with nonsense syllables and monosyllabic words, respectively, the essence of the verbal task they were asked to perform was, as in C1, letter-based: in all three experiments, subjects were asked to think about the alphabetical order or the number of consonants in the verbal units. This entailed that participants could potentially ignore the phonemic surroundings of the verbal element in question and concentrate on locating the consonants, in order to perform the task as quickly and accurately as possible. The reason these tasks were selected was to maintain comparability between tests in different languages; in order to accommodate a range of languages, however, the precision of the test regarding the nature of the interference tested was compromised.

The expected differences between the solfège and non-solfège groups were not located in the data analysed for this experiment. The only statistically significant difference between the two groups had, in fact, the opposite direction to the one expected in the hypotheses: in Experiment C3, where the subjects were asked to perform a parallel word-based task, solfège participant had produced significantly better Note span scores

than non-solfège subjects, indicating that, at least in terms of absolute comparison, the parallel verbal task did not affect their memorisation performance negatively. The possibility of solfège subjects not using solfège names to perform the memorisation task was considered very low, as the stimuli were presented only visually and subjects expected to demonstrate their recalled answer via solfège names; this allowed two possible explanations for the perceived effect:

1. Solfège subjects found the word-based task easier than their non-solfège counterparts. This could have happened for numerous reasons: the nature of the Greek language, the characteristics of Greek monosyllabic words and/or the education system, which made it easier for Greek participants to isolate consonants and count them within the given timeframe, were just a few of the possible reasons why Greek solfège participants could have presented a relative advantage in performing the distraction task.

2. The verbal task in C3, despite being presented in words, was principally a letter task; for this reason, non-solfège subjects, who were using letter names for the memorisation task, were negatively affected by the parallel performance of the task. Such an explanation, however, would be justified only if a similar negative effect had been found on the Note span scores in Experiment C1, in which participants had to perform a purely letter task. Since results in C1 showed no effect of solfège knowledge on Note span scores, it was not possible to attribute the significant effect in C3 to the design of the verbal task. It is true, however, that the ambiguity of the nature of verbal tasks in C2 and C3 – a letter task presented within nonsense syllables and a letter task presented within monosyllabic words, respectively – constituted a major restriction in the design, which prevented any generalisations in the conclusions. For C3, a semantic task based on words would help clarify the ambiguity; on the other hand, preserving the monosyllabic character of words would be impossible in such a task, mainly due to restrictions posed by the Greek language.

Regarding performance in the distraction task, analyses showed no difference between results produced by solfège and non-solfège subjects, in any of the three experiments; moreover, span scores and distraction task scores were found to have no significant correlation, apart from one case: solfège participants who scored highly in the word distraction task were found to be more likely to perform well in the memorisation task

in Experiment C3. This finding, however, was considered to be circumstantial and no further explanation was attempted to be given, as it was not consistent within the solfège group or any other group throughout the rest of Experiment C. The most notable finding regarding performance in the distraction tasks was the high percentage of solfège and non-solfège participants performing below the 70% correct threshold in Experiment C2; even after the removal of zero-scoring participants, 30% of the solfège and 43.8% of the non-solfège participants had got less than 70% of the answers in the distraction tasks correctly. This percentage is particularly important when seen in contrast with the equivalent percentages in the rest of the experiments in this unit; a comparison of the Total task correct scores across experiments will be discussed in the following chapter.

5.4. Experiment 3: General discussion and conclusions

5.4.0. Synopsis of Experiment 3

The aim of Experiment 3 was to illuminate aspects of the function of solfège note names within memory by comparing their effect with the effect of using letter names when performing two tasks in parallel; for this purpose, a series of computerised experiments, A, B and C was designed. Experiment A was designed to function as a basis for all further comparisons, as it combined two musical tasks: a memorisation task, for which the recall was executed via either solfège or letter note names, and a parallel pitch comparison task. In the B series of experiments, including B1, B2 and B3, the note memorisation task was replaced by a verbal memorisation task, whilst the parallel pitch comparison task remained the same as in Experiment A; in the C series of experiments, also including three subvariations, C1, C2 and C3, the memorisation task was musical, as in Experiment A, but the parallel task was verbal. Each one of the seven experiments (A, B1, B2, B3, C1, C2, C3) also contained variations in the number of parallel tasks performed in each experimental set (Experiments A and C) or in both the number of memorisation stimuli presented and the number of parallel tasks (Experiment B). The decision to include so many variations and subvariations of the basic experimental core stemmed from a lack of prior research in the field, combined with a lack of time to pilot each element of each experiment separately; piloting for this experiment aimed mainly to determine durations for stimulus presentations and the range of the number of items presented in each experimental set. As such, the whole of Experiment 3 was mainly exploratory in nature, aiming to provide a general framework for further research in the field; data analysis revealed not only the most prominent quantitative indexes but also certain characteristics of the dataset which required the adoption of a different methodological approach in order to be tested thoroughly. On the other hand, an inherent drawback of this type of non-restrictive, broadly-conceived design was that it rendered the generalisation of any conclusions drawn implausible; all findings from data analysis have to be considered within the framework and limitations of the present research.

The task which the participants were asked to perform in parallel with the memorisation task could either be a pitch comparison, as in experiments A and B, or a verbal task; the importance of performing the parallel task as accurately as possible was stressed repeatedly to the participants before the completion of each experiment. In order to ensure that participants were not giving priority to the memorisation task, neglecting the parallel task, a 70% threshold was established as the minimum amount of correct answers in the parallel task, over which participants were considered highly likely to have invested sufficient effort in the performance of the task. Subjects scoring below 70% correctly in the distraction task were not automatically classed as subjects who had not attempted to perform the task appropriately, but rather as displaying considerable difficulty in performing the two tasks simultaneously; for this reason, data from these subjects were excluded in some of the analyses. On the other hand, data from subjects who had produced zero correct answers in the distraction task were excluded from all analyses, as these individuals were thought to have effectively performed only the memorisation task. Many subjects had reported, after the completion of the experimental procedure, that they had found themselves inadvertently pressing the wrong button on the computer keyboard in responding to the distraction stimulus; as, according to self-reports, none of the subjects were dyslexic or suffering from other learning disabilities, and given the extensive amount of practice trials taking place at the beginning of each experiment, such 'reflex errors' were considered to have been distributed evenly amongst participants, so their effects had been balanced. It is important to remember at this point that correlation between distraction task scores and span scores was found significant in very few, sporadic cases: for non-solfège participants the relationship was significant in Experiments A and B1, while for solfège participants the relationship was significant only in Experiment B3. The direction of the correlation, which was positive for both groups, demonstrated that, even in these cases, low performance in the distraction task was not a result of favouring one task over the other, but rather a reflection of a more general difficulty in performing both tasks simultaneously; this resulted in consistently high or consistently low scores in both tasks.

In summary, data analysis in each of the three experiment groups, A, B and C, showed that:

1. Non-solfège participants who performed well in the parallel music task were more likely to perform well in the memorisation task as well, when the memorisation stimuli were notes or letters; when the memorisation stimuli were nonsense syllables or words, no correlation was found between performances in the two parallel tasks for the solfège group.

2. Solfège participants who performed well in the music memorisation task were more likely to perform well in the parallel verbal task, when the stimuli were monosyllabic words and the task was to determine the difference in the amount of consonants in each word pair; no correlation was found between performance in the music memorisation task and performance in the verbal task when the latter required the alphabetic classification of letters, either presented in isolation or within nonsense syllables.

3. Solfège and non-solfège participants did not show any statistically significant difference in their performance of the music memorisation task when the parallel task included the classification of pitch heights or alphabetical order.

4. Non-solfège participants appeared to yield significantly lower scores in their performance of the music memorisation task when the parallel task included the arithmetic comparison of consonants within word pairs.

5. Solfège participants appeared to yield significantly lower scores in the verbal memorisation task compared to non-solfège participants, when the memorisation stimuli were nonsense syllables; the difference between solfège and non-solfège performances when stimuli were words was found borderline non-significant (p = .058). No significant difference was found between solfège and non-solfège performances when verbal memorisation stimuli consisted of letters.

5.4.1. Comparison of performance in experiments A and C

Experiment A and experiments in the C series (C1, C2, C3) shared the same pitch memorisation tasks but had different distraction tasks. Within each one of these experiments separately, no difference was found between solfège and non-solfège performances, in either the memorisation or the distraction task, apart from Experiment C3: in C3, when the distraction task consisted of the arithmetical comparison of

consonant numbers in pairs of words, solfège subjects were found to be more likely to yield higher Note span scores.

The general design adopted in Experiment 3, according to which each participant was tested on a pseudo-random selection of three or four experiments out of the seven available, had secured an adequate number of participants in each experiment separately: the minimum number of participants in an experiment was 41, for experiments C1, C2 and C3 and the maximum 55, for Experiment A. This approach, however, did not control for the number of subjects participating in combinations of experiments; this restriction, along with the fact that most score distributions were non-normal, hindered repeated-measures analysis of data from participants who had been tested across several experimental conditions. Moreover, although sample sizes justified conducting mixed ANOVA in each pair of tests, for example on data from A and C1, A and C2, A and C3 and so on, using Note span as the within-subjects variable and solfège group as the between-subjects variable, the inequality of solfège and non-solfège group sizes, combined with the lack of within-group normality and the violation of the homogeneity of variance assumption in the vast majority of the cases, would have rendered the resulting F statistics extremely unreliable.

For these reasons, Δ scores were used again (see Chapter 3.3), representing the difference between Span scores in two experiments: for example, if Note span achieved in Experiment C1 was subtracted from Note span achieved in Experiment A, the resulting Δ score (Δ_{AC1}) would indicate whether or not the subject had performed better in one of the two tasks; a negative value would indicate that Note span had been higher in C1, a positive value would indicate that Note span had been higher in A and a value of zero would indicate that the subject had remembered the same total amount of notes in both experiments.

Δ scores were calculated for the following combinations of Note span scores:

Δ_{AC1} = Note span A - Note span C1 (N = 19)

Δ_{AC2} = Note span A - Note span C2 (N = 22)

Δ_{AC3} = Note span A - Note span C3 (N = 22)

Δ_{C31} = Note span C3 - Note span C1 (N = 16)

Δ_{C32} = Note span C3 - Note span C2 (N = 19)

Δ_{C12} = Note span C1 - Note span C2 (N = 14)

Although sample sizes were still insufficient, each one of these Δ scores was entered as an outcome variable in simple regression analysis using bootstrap samples, with solfège group as the predictor variable with two levels: solfège and non-solfège; all regression results were found non-significant, suggesting that the change in Note span scores did not depend on whether or not the participant had used solfège in recall.[7]

Mean scores for each group (solfège, non solfège and taught solfège) in each experiment were also plotted, in order to acquire a more spherical view of the data gathered from all experiments; the lines in Figure 5.18 represent Mean scores achieved by each group in Experiments A, C1, C2 and C3:

Mean Note span scores (full sample)
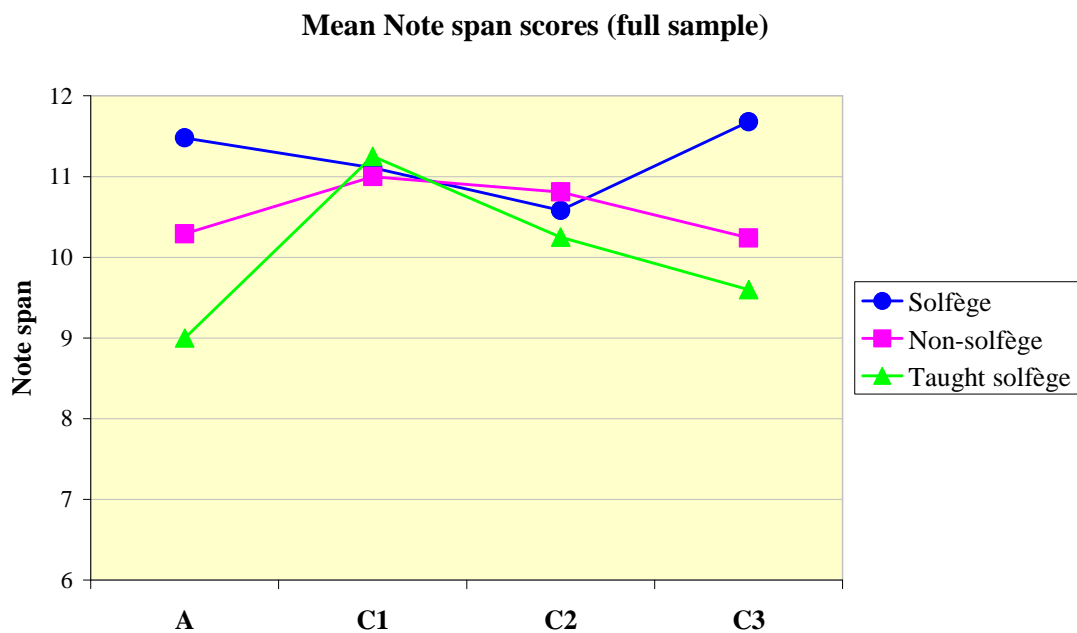
Figure 5.18: Mean Note span scores achieved by participants in the experiment. Note that the above graph shows full sample means, rather than means for participants who completed all four experiments; therefore sample sizes are equal to sample sizes in each experiment separately, minus all zero-scoring participants in the distraction task.

---

[7] Simple regression results for each analysis separately can be found in Appendix 5.

The lines in Figure 5.18 agree with non-significant regression findings in all cases but those which involved data from Experiment C3: mean scores for the solfège group in this experiment appear higher compared to the other experiments, especially the lower-scoring C2. Solfège subjects appear to have performed better than the other groups in C3, a finding consistent with regression analysis results in C3 (see section 5.3.3.6); they also appear to have performed better in C3 as a group compared to performance in the other Experiments. Another interesting point emerging from Figure 5.18 was that the taught solfège group means follow the same pattern as the non-solfège group means rather than that of the solfège group; this is more apparent in the difference in means between Experiments A, C1 and C3, as the mean Note span of all three groups appears to have dropped in C2 compared to C1. The same conclusions can be drawn from the line graph depicting the amount of subjects in each experiment scoring less than 70% (9 notes) in Note span, as shown in Figure 5.19:

Note span scores below 70%
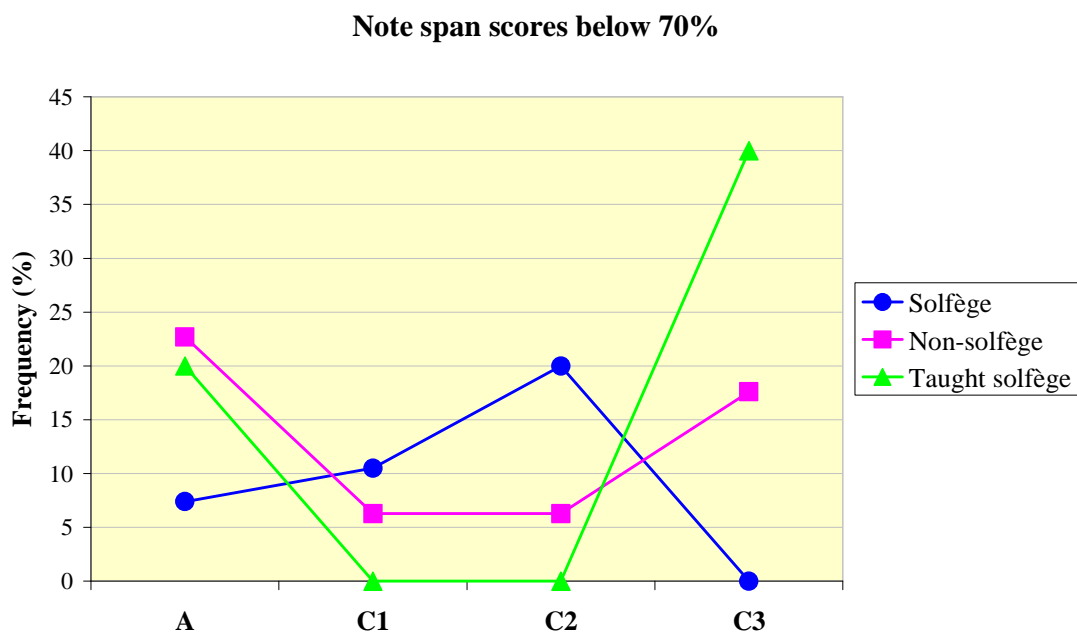
Figure 5.19: Percentage frequency of participants remembering less than 9 notes out of 12 in experiments A, C1, C2 and C3.

Figure 5.19 indicates that, as suggested by Figure 5.18 and the results of regression analysis in 5.3.3.6, solfège participants performed better in C3 compared both to the other experiments and to the other groups, as there were no solfège participants that

remembered less than 9 notes in C3, while the non-solfège group had its biggest population with a Note span of less than 9 notes in the same experiment. The exact opposite appeared to happen in Experiment C2, in which the non-solfège group had zero participants scoring below 70% in Note span, while at the same time the solfège group had its highest percentage of below-70% Note span subjects in C2; it is also interesting to note that zero non-solfège participants had a Note span of less than 9 in Experiment C1 as well. Taught solfège group appears again to yield a pattern closer to the non-solfège group than to the solfège group. These findings suggest that, although solfège knowledge groups yielded relatively similar Note span scores, in absolute terms, within the same experiment, there were differences between groups in the way Note span fluctuated across experiments; the fact that Δ scores did not corroborate this finding could be attributed to the extremely small sample sizes available for comparisons.

It is also interesting to see how scores in the distraction task changed across experiments; since the 70% threshold was adopted for the distraction task throughout Experiment 3, it is useful to see the percentage of solfège, non-solfège and taught solfège participants scoring below the 70% threshold in Experiments A and C (Figure 5.20):

Distraction task below 70%
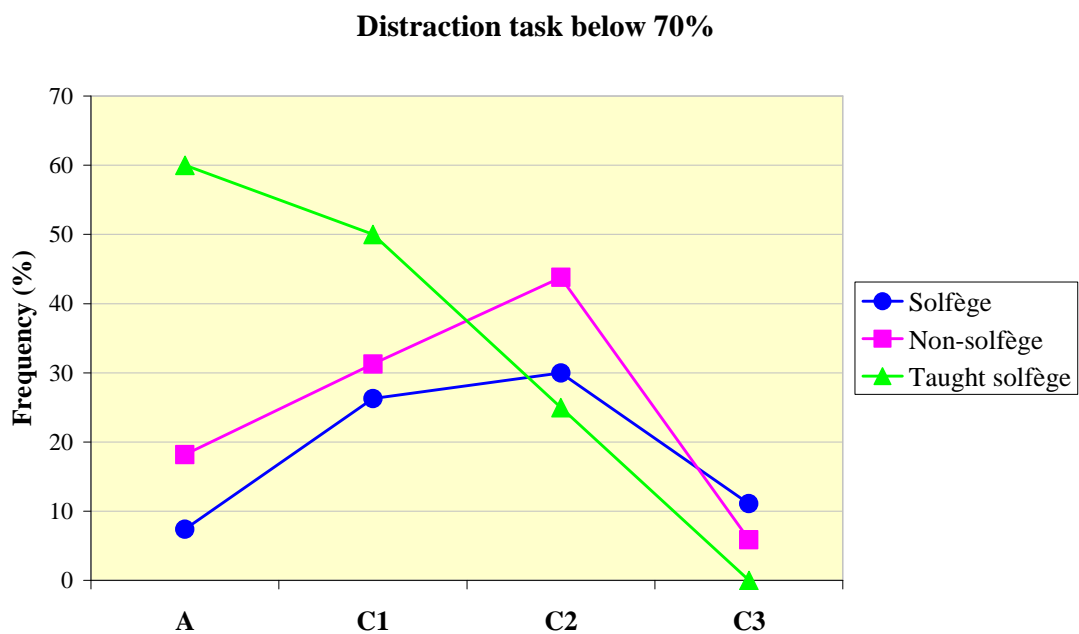
Figure 5.20: Percentage frequency of participants scoring below 70% correctly in the distraction task in each experiment.

Figure 5.20 indicates that both the solfège and the non-solfège groups presented a similar pattern of results regarding the distribution of participants in high (70% or over) and low (below 70%) scorers in the distraction task. It is important at this point to note that, in each of the four experiments, participants had to perform a different distraction task; in Experiment A the distraction task was a pitch comparison task, in C1 participants had to determine alphabetical priority in pairs of letters, in C2 participants had to determine alphabetical priority of letters in pairs of nonsense syllables and in C3 participants had to compare the number of consonants in pairs of words. Despite the differences in the nature of distraction tasks, however, subjects appeared to produce similar patterns of high and low scorers in each experiment; the fact that for both the solfège and the non-solfège groups C2 had the most low scorers and C3 the least, suggested that performance in the distraction task was affected by factors like the difficulty of the task overall and perhaps the difficulty of the simultaneous performance of the two tasks. This called for an alteration in the experimental design, so that subjects would be 'forced' to perform over 70% in the distraction task; the Unsworth et al (2005) Aospan experiment, on which all experiments in this section were modeled, provides a possible solution to this problem (Unsworth et al, 2005).

Apart from pictorial representation of the data, the observation of the actual percentage values showing the frequency of participants scoring below 70% in the distraction task can also lead to useful remarks: Table 5.26 displays these frequencies for the solfège and the non-solfège groups; taught solfège data were omitted because of the limited sample size.

Occurrence of Total task correct scores below 70%

	A	C1	C2	C3
Solfège	7.4%	26.3%	30%	11.1%
Non-solfège	18.2%	31.3%	43.8%	5.9%

Table 5.26: Percentage frequency of participants scoring below 70% in the distraction task in each experiment.

The values in Table 5.26 reveal that Experiments C1 and C2 had produced a considerable amount of low-scoring participants; with a minimum of 26.3% and a maximum of 43.8%, the tasks including alphabetical ordering in C1 and C2 appeared to pose serious difficulties to participants, compared both to the music distraction tasks of Experiments A and B and to the consonant-counting task of C3. This finding, too, suggested the need to alter the presentation of the distraction tasks in order to avoid such poignant fluctuations in the data; relatively stable scores in the distraction tasks would help towards a clearer interpretation of possible effects of the distraction task on the memorisation and recall tasks.

5.4.2. Comparison of performance in experiments A and B

Experiments A and B had the distraction task in common and differed in the nature of the memorisation tasks: in Experiment A subjects were asked to memorise pitch sequences, in B1 letter sequences, in B2 nonsense syllable sequences and in B3 monosyllabic word sequences. Analysis of experimental data yielded a significant difference between the solfège and the non-solfège group in Syllable span scores in B2 (section 5.2.2.6) and a borderline non-significant difference between the two groups in Word span scores (section 5.2.3.6).

Δ scores were also calculated for all combinations of distraction scores; in order to distinguish them from Δ span scores, Δ distraction task scores are denoted as ΔD scores:

$\mathbf{\Delta D_{AB1}}$ = Total task correct A - Total task correct B1 (N = 25)

$\mathbf{\Delta D_{AB2}}$ = Total task correct A - Total task correct B2 (N = 25)

$\mathbf{\Delta D_{AB3}}$ = Total task correct A - Total task correct B3 (N = 25)

$\mathbf{\Delta D_{B31}}$ = Total task correct B3 - Total task correct B1 (N = 17)

$\mathbf{\Delta D_{B32}}$ = Total task correct B3 - Total task correct B2 (N = 18)

$\mathbf{\Delta D_{B12}}$ = Total task correct B1 - Total task correct B2 (N = 18)

Note that Δ span scores between Experiments A and B were not calculated, as they measured different entities; Experiment A span scores were note span scores and

Experiment B span scores were letter, syllable, or word span scores. All scores in the above calculations for ΔD scores were used as percentage values.

Although sample sizes were not sufficient, each one of the ΔD scores was entered as an outcome variable in simple regression analysis using bootstrap samples, with solfège group as the predictor variable with two levels: solfège and non-solfège. All regression results were found non-significant,[8] with the exception of ΔD_{B32}: regression results for ΔD_{B32} are presented in Table 5.27:

ΔD_{B32} scores

	B	SE B^a	β
Constant	- 3.97	3.49	
Solfège knowledge	10.46	- 0.11	- .53**

$R^2 = .28, p < .05$

** $p < .05$, *s*

a. Bootstrapped values. Significance for β is also based on bootstrap samples.

Table 5.27: Simple regression results for ΔD_{B32} scores.

The value of R^2 indicates that solfège knowledge had a significant effect on the change of Total task correct scores between Experiments B2 and B3. The value of beta shows that participants belonging in the solfège group were more likely to perform better in the music distraction task in B3 than in that in B2, or, conversely, non-solfège participants were more likely to perform worse in the music distraction task in B3 than in that of B2. Since the music distraction tasks were very similar in the two experiments, there were three plausible explanations for this effect:

1. The difficulty solfège participants had met in the memorisation task in B2 (see section 5.2.2.6) had had a negative effect on their performance in the distraction task as well.

2. Non-solfège participants had performed worse in the distraction task of B3 than in that of B2 due to factors related to the performance of the word or the syllable memorisation task.

---

[8] Simple regression results for each analysis separately can be found in Appendix 5.

3. A combination of the above.

A graphical representation of Total task correct scores across experiments helped on the interpretation of regression results, as shown in Figure 5.21:

Total task correct below 70%
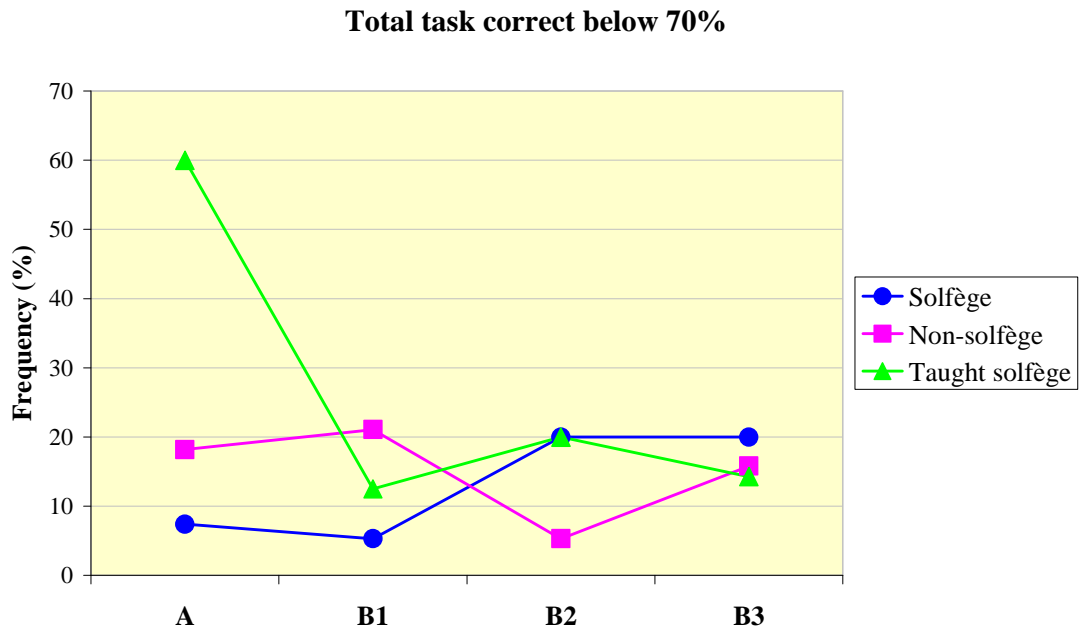
Figure 5.21: Percentage frequency of participants scoring below 70% correctly in the distraction task in each experiment.

Figure 5.21 indicates that, in the solfège group, there was a consistent, relatively high amount (20%) of participants yielding low scores in the distraction tasks; on the other hand, the non-solfège group appeared to have a relatively small amount of low scoring participants in B2, while their amount increased in B3. Moreover, mean scores for the distraction task in the two experiments were very close for high-scoring subjects in the solfège and the non-solfège group (see Table 5.28).These findings, along with significant regression results for Syllable span scores in Experiment B2 suggested that the differences between the two groups were a result of higher distraction task scores in B2 by the non-solfège subjects, rather than that of excellent performance in B3 solfège participants. As mentioned previously, distraction tasks were the same in all B experiments; the reasons behind this change in performance by non-solfège subjects

should therefore probably be traced in the nature of the syllable memorisation task in B2 and perhaps the different linguistic traits characterising the two languages.

Mean Total task correct scores of over-70% participants

	B2	B3
Solfège	87.44	89.41
Non-solfège	88.44	89.33

Table 5.28: Mean scores achieved in the distraction task by participants who had performed over the 70% threshold.

Lastly, letter, syllable and word span scores in the separate experiments had only revealed a significant difference between the solfège and the non-solfège groups in Experiment B2, in which solfège participants appeared to yield significantly lower syllable span scores than non-solfège participants. Although different entities, letters, syllables and words all constituted the sub-elements of the verbal component of the experimental design; for this reason, it was considered plausible to compare letter, syllable and word span in this context. The following Δ scores were calculated for all combinations of span scores from experiments in group B:

Δ_{B31} = Word span B3 - Letter span B1 (N = 17)

Δ_{B32} = Word span B3 - Syllable span B2 (N = 18)

Δ_{B12} = Letter span B1 - Syllable B2 (N = 18)

Simple regression analysis was conducted using each one of the above Δ scores as the outcome variable and solfège group as the predictor, with two levels: solfège and non-solfège; bootstrap samples were used for the analysis. All regression results were found non-significant at p > .05;[9] due to the restrictive number of the samples, however, Span scores were also plotted in order to acquire a more spherical view of the data. Figure 5.22 shows the change in the percentage of participants in each experiment who scored below 70% in the memorisation task:

---

[9] All regression results can be found in Appendix 5.

Span below 70%
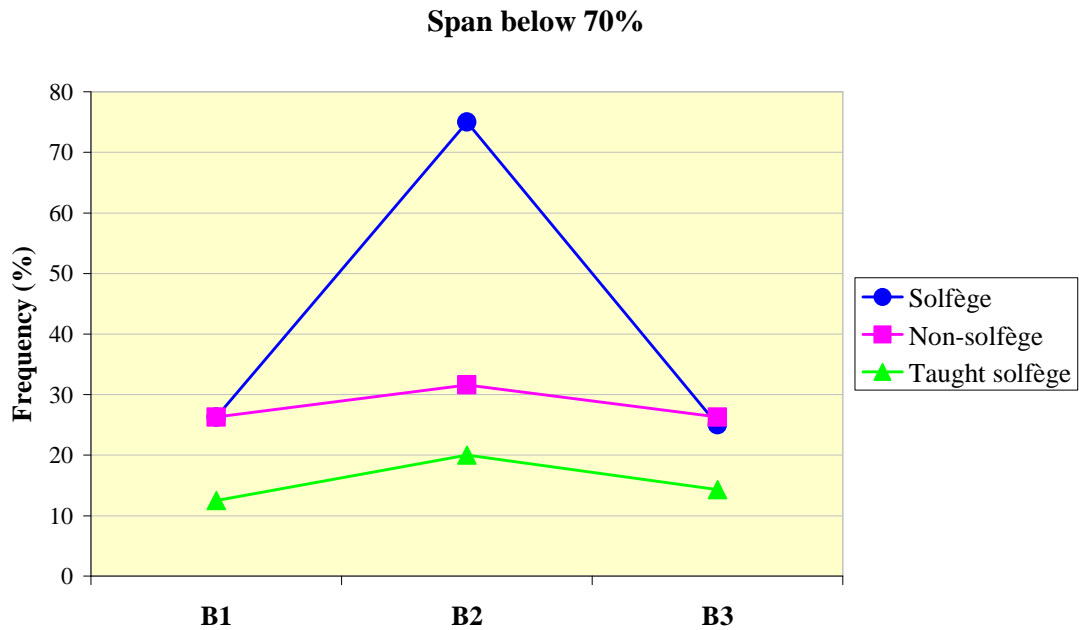
Figure 5.22: Percentage frequency of participants scoring below 70% correctly in the memorisation task in each experiment.

Figure 5.22 does not reveal any new evidence regarding memorisation performance in the verbal tasks; the most prominent change in scores was the one incurred in B2 by solfège participants, 75% of whom failed to remember more than 70% of the nonsense syllables in the memorisation task. It is important to note that solfège and non-solfège span scores in B1 and B2 were almost identical, suggesting that there was little possibility that the significant change in B2 was a result of a difference in language skills; on the other hand, given the difference between the linguistic attributes of phonemes and lexical morphemes, specified testing is required in order to interpret the significant difference between the two groups. The hypothesis in B2 posited that the difference between the groups would result from the interference of the pitch comparison task using solfège syllables with the memorisation of nonsense syllables; one way to provide further support for this hypothesis would be to investigate whether the significant result would be replicated in a condition where the distraction task would not involve pitches but digits or images: the replication of results would suggest that low performance in the memorisation task stemmed from a general difficulty of solfège

subjects with memorising nonsense syllables, rather than from the interference caused by the pitch comparison task.

5.4.3. Summary

Experiment 3 constituted a first step in the investigation of the nature and function of solfège note names in human memory and cognition; although further research is required in order to extract more definitive results regarding the musical and linguistic characteristics of solfège syllables, certain conclusions emerged from the analysis, which can serve as points of reference for future research projects:

1. Solfège and non-solfège participants did not appear, in most cases, to produce significantly different patterns of results in absolute terms, compared within each experiment separately; the plotting of mean scores, however, revealed that there might be distinct underlying patterns for each group across experiments. In order to test this efficiently, a different sampling approach would be required: instead of recruiting participants so as to maximise the number of subjects per experiment, a matrix of meaningful experiment combinations should be constructed in order to maximise the number of participants tested in each combination.

2. On the basis of the longitudinal examination proposed in the previous point, both verbal and musical distraction tasks should be revised so that they include more target-specific stimuli, even if these are not exactly identical across languages. An example from the verbal tasks would be changing the word task from a consonant-counting task presented within the framework of words, to a semantic distinction task, in which subjects would be required to determine whether or not, for example, two words belong in the same category (e.g. dog – cat, dog – chair), or to pick the irrelevant word from a set of three (e.g. dog – cat – straw).

3. Despite the piloting that had taken place in order to define stimulus durations for the experiments, mean Note span scores in experiments A and C were very close to maximum (minimum Note span = 84.31% in C3 by the non-solfège group).

This finding revealed the need for further piloting and possible extension of the number of pitch stimuli, in order to avoid ceiling effects.

4. In order to corroborate and interpret the significant difference between the solfège and non-solfège groups in B2 syllable spans, more extensive investigation of solfège participants' syllable span would be required. A possible way to test the validity of the recorded effect and its relationship with the use of solfège names in the distraction task, it would be useful to compare Syllable span scores in experiments with different distraction tasks; if the effect did not persist when the nature of the distraction changed, the effect could more definitively be attributed to the interference of solfège names with memory for phonemes.

Chapter 6

General discussion
and
conclusions

6.1. General remarks and conclusion

The present research constituted an ambitious first step towards a contemporary approach of solfège as a pitch labeling system. As the largest part of the extant literature on solfège mainly addressed its pedagogical value and implications, this study aimed to illuminate the conceptual and cognitive implications of using solfège syllables for labeling pitches, adopting an empirical approach. In order to achieve this, the use of solfège syllables was compared to the use of letter names throughout the study; the selection of letter names was made purely on the basis of their widespread use across countries and cultures and on the common characteristics they share with other music-reading systems, such as the use of numbers, without implying any form of duality in the range of music reading systems available. The research questions, which formed the basis and determined the structure of the thesis, were the following:

1. Does the early learning and use of solfège note names affect music memorisation performance in adulthood?

2. Does the early learning and use of solfège note names affect long-term retention of memorised music in adulthood?

3. Does the use of solfège note names affect music memorisation performance when solfège has been learnt during adulthood?

4. Does the use of solfège note names affect long-term retention of memorised music when solfège has been learnt during adulthood?

5. Is solfège different from other note-reading systems, in terms of the interaction taking place between note names and linguistic elements?

The aims determined by the research questions and conclusions emerging from their empirical examination within the framework of the present study are going to be addressed separately in the subsequent sections.

6.1.1. Aim 1: To examine the effects of solfège use on a music memorisation task performed by adult musicians.

Findings from Experiment 1 did not support hypotheses predicting that solfège subjects would yield different performance scores than non-solfège subjects: solfège knowledge and solfège use did not appear to affect performance scores. On the other hand, analyses showed that performance scores in the tonal and the atonal pieces were affected by the instrument performed, with wind and string instrument players appearing to be the most competent memorisers compared to singers and pianists. This study is the first to make a direct comparison of memorisation performance between musicians performing on different instruments; the significant differences displayed between instrumental groups offer a stimulating topic for further research in the field. A partial explanation for the differences in memorisation performance by different instruments in this study can be possibly located on the time usually required to learn and memorise a piece, as the strict timeframe of the present study may have caused some expert memorisers from the singer and the pianist groups to perform worse than if they had more time to learn the piece.

Other factors affecting the memorised performance of the tonal pieces appeared to be the frequency with which subjects used to memorise music and the strategy adopted to memorise the piece. High-frequency memorisers appeared to have an advantage over students who were not used to memorising music for performance; furthermore, subjects who used the 'analytical' strategy, including studying the piece mentally, without performing the music on the instrument, produced better performances than subjects using other strategies. The importance of the analytical strategy has been repeatedly supported in literature (Chaffin and Imreh, 1997; Hallam, 1997; Williamon, 2002); it is interesting to note, however, that participants' responses regarding the memorisation techniques used as well as the use of solfège did not always correspond to their actions as recorded in the practice and recall sessions. A similar finding has been reported by Ginsborg (2002), in a study comparing memorisation techniques of singers with varying levels of expertise; the discrepancy in the present study was also followed by a high percentage (80.65%) of participants reporting that they had never been taught how to memorise music. This could indicate a gap between theoretical knowledge and

practical application: although most music students appear aware, either due to intuition or as a result of academic teaching, of the advantages of analytical memorisation strategies, they display difficulty in applying this theoretical knowledge in their music practice; this could be a call for instrumental teachers to put separate, specific emphasis on memorisation as a key aspect of excellent music performance.

Performance of the atonal pieces appeared affected by instrument type but not by memorisation frequency; the memorising strategy that led to better scores in the performance of the atonal piece was the division of the piece in smaller sections, also known as chunking. The non-significance of the level of familiarity with atonal music contradicted findings from studies investigating the importance of familiarity of the stimulus in music memorisation, compared to other factors: the level of expertise on the instrument has been found to be affecting performance of tonal but not atonal music (Knecht, 2003) and similar findings have emerged from studies examining differences between memorisation of culturally familiar or unfamiliar music (Demorest et al, 2008), and tonal or modal music (Oura and Hatano, 2004). These studies have shown that lack of familiarity with the musical stimulus eliminates differences in memorisation performance caused by differences in expertise; in the present study, however, the instrument performed appeared to be a more definitive factor than familiarity with the stimulus or memorisation expertise, to the degree this can be expressed by memorisation frequency. The apparent overriding importance of the type of instrument performed in the memorisation performance of both tonal and atonal music deserves to be investigated further in a dedicated study; in order to eliminate time limit as a potential factor that might be affecting results, by placing a disadvantage on competent but slow memorisers, the differences between instrument types in memorisation performance would be better investigated via an observational study which would fully record the adopted approach during a more extended period of time; this would allow the extraction of safer conclusions regarding the reasons behind these differences.

6.1.2. Aim 2: To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians.

The number of missing cases in the 3rd, after 24-hours, performance of the task did not allow rigorous analysis to be conducted on the data; comparison of Δ scores, however, revealed big differences in the numbers of solfège and non-solfège participants performing better in the 3rd, compared to the 1st, performance of the tonal piece: between 15.2–21.5% more solfège participants than non-solfège participants improved in the 3rd performance.[1] Although it could be argued that this difference could be a result of lower solfège scores in the 1st performance, rather than higher scores in the 3rd, analysis results showing that solfège and non-solfège performance scores were similar renders this possibility highly unlikely. The magnitude of the frequency differences in Δ scores suggest that there could very possibly be a genuine positive effect of solfège knowledge and use in the long-term retention of pitch sequences; this could be investigated further through a longitudinal study of solfège and non-solfège musicians, so that any restrictions regarding the approach adopted between experimental sessions would be neutralised between groups.

6.1.3. Aim 3: To examine the effects of solfège use on a music memorisation task performed by adult musicians who have been taught solfège recently.

The data gathered in Experiment 2 violated several assumptions required for conducting parametric tests, therefore Δ scores were used in the analyses. This precluded separate analysis and comparison of each of the three performance scores produced by each participant before and after they had learnt solfège; Δ scores, however, revealed that there was a significant effect of solfège on the change of score between the 1st and the 3rd performance of the tonal piece, with subjects who had used solfège improving more than controls. Although this pertained more to Aim no. 4, regarding the long-term retention of music, it still constituted a significant effect of solfège use on

---

[1] The values of 15.21% and 20.5% represent the percentage of subjects who performed better in the 3rd performance using explicit and implicit solfège knowledge, respectively.

memorisation, albeit in relative rather than absolute terms. Some of the most important findings in Experiment 2 were related to the process of learning solfège in adulthood; the reception of solfège classes was extremely positive and students gave enthusiastic feedback regarding the contribution of solfège in honing their aural skills in general. Moreover, most subjects who participated in the solfège programme as part of the experiment were able to use solfège as instructed for the purposes of the experiment; as fixed-do solfège is almost exclusively taught to novice musicians at a young age, the present study was useful in showing that efficient solfège learning is both attainable and beneficial during adulthood, even for students who are already proficient in music reading.

6.1.4. Aim 4: To examine the effects of solfège use on the long-term retention of a pitch sequence by adult musicians who have been taught solfège recently.

The hypotheses relative to this aim were upheld in the analysis for Experiment 2; as mentioned in the previous section, analysis of Δ scores revealed an increased likelihood for solfège-using participants to improve their performance between the 1st and 3rd performance of the tonal piece. As the study was principally based on observation, it was not possible to determine the exact cause of this effect; further testing would be required in order to investigate whether the apparent long-term advantage offered by solfège in music memory has its roots in specific characteristics of solfège syllables, which make them more durable over longer time-spans, or in the structure of the process followed in learning and memorising a piece using solfège.

6.1.5. Aim 5: To examine the differences between subjects using solfège syllables and subjects using letter names, regarding the extent and nature of interference between verbal and musical tasks being performed simultaneously.

Experiment 3 was useful in laying the groundwork for the construction of a paradigm which can be used to illuminate the function of solfège syllables as cognitive labels within perception. Despite the inevitable restrictions in the design of the experiments, several useful findings emerged from analysis, especially from experiments investigating the interference of music tasks in the memorisation of verbal elements. Analysis revealed a significantly lower performance in solfège subjects when the memorisation stimuli were nonsense syllables and a borderline non-significant performance when memorisation stimuli were words, whilst no difference was found when stimuli consisted of letters. Despite the aforementioned methodological restrictions, these results can be explained by the model proposed in Chapter 1.2 for solfège syllables: lower performance scores in the memorisation of nonsense syllables could indicate the dissociation of solfège syllables from the syntactical category of words, despite their clear-cut, exclusive correspondence with specific elements of reality – notes. On the other hand, the lack of significantly lower performance scores by non-solfège subjects when the memorisation stimuli were letters could indicate the proposed absence of such strict correspondence between signifier and signified in letter names for notes: as letters have principally a different use than signifying pitches, musicians who are taught music via letter names do not form automatic cognitive representations of notes linked to their letter names; as a consequence, music does not constitute interference in a letter-memorisation task. Further testing is necessary in order to corroborate these theories and define the exact mechanisms underlying the interaction between solfège syllables, verbal elements and letter names; using the present study as a starting point, numerous experiments can be conducted in order to investigate these effects further.

On the other hand, the significantly lower scores in the music memorisation task by the non-solfège subjects was ambiguous, as the distraction task included a letter-based task, albeit presented within words. Once more, the design of the experiment and the need to

produce comparable tasks for both the solfège and the non-solfège groups restricted the attempts to interpretation; the duplication of significant results in conditions where the subjects would be required to perform a purely word-based or a purely letter-based task, would help determine the cause of the effect.

6.2. Limitations

The present study set out to explore the role of solfège in the context of music memorisation practice, performed by highly-skilled musicians. More specifically, Experiment 1 examined memorised performances by students from ten different countries, divided in five groups (Chapter 3.2.2.1); although all the stimuli used belonged in the Western classical music tradition and were administered to students playing instruments used in, but not limited to, this tradition, the participants' music education background varied greatly according to their country of origin. Moreover, many of the participants were not primarily performers of classical music, but instead jazz, popular, or, in fewer cases, Eastern traditional musicians. Although time and resources restrictions did not allow for a more systematic sampling of participants across a variety of music cultures, the present study was cross-cultural to the extent that different pedagogical systems and different listening and performance preferences can be known to form cultural variants. On the other hand, the cross-cultural factor in Experiment 1 did not include musicians belonging exclusively to a music culture other than the Western; although there are parallels between solfège and other notation systems, as, for example, the Indian sargam, an empirical investigation of the similarities between solfège and a system belonging to an entirely different musical tradition would require an entirely different design to that of the present study in terms of stimuli, participant selection and focus, which could form the basis of a future research project.

Apart from restrictions in participant selection, mainly imposed by practical factors, another limitation of the study concerned the nature of the stimuli used in Experiments 1 and 2 (Appendix 2): in order to ensure participants' focus on the memorisation process, the musical stimuli chosen for these experiments were only average in terms of their technical level. The decision to use the chosen stimuli emanated from the need to

limit the number of independent variables affecting memorisation performance: in addition to the screening of participants according to their expertise on their instrument or voice, the selection of relatively easy musical excerpts ensured that differences in memorisation performance would not be a result of differences in the technical command of the piece. This did not imply that the pieces were simplistic in terms of structure and thematic content; on the contrary, they presented memorisation challenges for most of the participants in the study (Chapters 3.2.1.1 and 3.2.1.2). Nevertheless, the uniformity in the technical level of the pieces meant that findings in Experiments 1 and 2 are necessarily confined to a given level of musical complexity and cannot be generalised to memorising harmony or more complex polyphonic structures. As the present research was based on the study of solfège as practised through overt or covert singing and given the time restrictions in the study, musical pieces containing more than a single monophonic line could not be examined using the same methodological tools. Automatically, this also restricted the range of memorisation techniques studied: even in the case of pianists, whose stimuli consisted of simple 2-part polyphony in order to reflect the average technical challenge presented in canonic piano pieces, the methodology used in the study was not fit to monitor memorisation techniques other than solfège; all other strategies referred to in the analysis (Chapters 3.3 and 3.4) are based on meta-musical and meta-linguistic information derived from questionnaires and recording transcriptions.

The particularities of piano music as opposed to vocal, string and wind music had an additional implication for data analysis, as pianists were found to yield significantly lower memorised performance scores than other instrumentalists (Chapter 3.3). This was partly attributed to the fact that, on average, pianists had 92% more notes to learn than other participants; as an inherent methodological trait of the study led to the formulation of an entirely different hypothesis than the one the experiment had set out to explore (Chapter 3.4), the whole framework of hypothesis testing was revisited post-analysis. As the present study was the first to explore solfège using a rigorous quantitative approach, the original hypotheses were dictated by intuition and anecdotal, empirical evidence and were subsequently disproved on many occasions. Although the formation of hypotheses had been necessary in order for them to function as 'reference points', around which all testing revolved, many secondary hypotheses and parameters were revealed and tested during the experimental and the analytical process. As such,

hypotheses in Experiments 1 and 2 can be viewed, in retrospect, as starting points for a rigorous study of solfège, rather than an end in themselves: instead of producing absolute measurements of solfège contribution in the memorisation process, in most cases data analysis from the two experiments illuminated certain more general aspects of the incorporation of solfège in music memorisation.

The complexities of incorporating specific memorisation techniques into practice, combined with the methodological approach adopted in the study, resulted in one of the major ambiguities of the present research: although participants were verbally instructed to use certain techniques and also produced verbal reports of these techniques, it was not possible to trace exactly what took place during the 10-minute practice sessions. The methodology of choice for monitoring cognitive processes in real time is brain imaging: fMRI, MEG or EEG and PET are some of the most popular methods used by neuroscientists in order to detect changes in neural activity whilst performing given tasks. The adoption of any of these methods at this stage of solfège research, however, presented two basic problems:

1. The aforementioned methods are quite costly, as they require the use of highly sophisticated equipment and specialised staff. For this reason, it is highly unusual to adopt such methods before the analysis of preliminary research data, such as those presented in the previous chapters, which provide the necessary grounding for the formulation of specific hypotheses to be tested by neuroimaging methods.

2. All neuroimaging scans have to be performed in laboratory settings and participating subjects need to be physically connected to scientific equipment, ranging from the simple attachment of electrodes to subjects' heads in the case of EEG to the enclosing of the whole body in an MRI tube. In Chapter 2.0, the choice of a 'naturalistic' methodology for the present study was argued over a lab-based approach; philosophical and epistemological considerations aside, however, it would be impractical to design a brain imaging study which would monitor solfège use in different categories of instrumentalists at this early stage of research. Such methods are not suitable for detecting general, large-scale phenomena, as the ones tackled in the present study, but are rather reserved to locating the effects of cognitive activity during specified periods of time. For this reason, the lack of previous research in the field combined with highly

specialised neuroimaging data from subjects performing musical tasks under extremely unfamiliar circumstances would almost certainly yield obscure and inconclusive results.

Nevertheless, data from Experiment 2 could form the basis of further neuroimaging research in the future; the apparent change in memorisation performance taking place after solfège instruction (Chapter 4.4.2) can be further explored by monitoring neural activity during the production of music using and not using solfège syllables. Moreover, the significant results in Experiment 3 (Chapter 5.4.3) would be best followed up by examining the similarities and differences in the clinical image of the brain processing linguistic, musical and solfège data; such a follow-up study would also help clarify the ambiguity created by contradictory data in Experiment B2 and C3 (Chapter 5.4).

6.3. Research implications and directions for further study

The implications of the findings of the present study are mainly pedagogical, as solfège is primarily a method of teaching and learning music. The intention of this research was not to re-ignite an obsolete debate around the merits of solfège, its potential advantages over other music learning systems or even draw comparisons between the fixed- and movable-do systems; rather, it aimed to offer an alternative perspective on a traditional approach, away from the rigid, obsolete systems of the past, aiming to show ways to integrate solfège in a multi-dimensional, mixed-methods approach to music teaching, from which musicians will benefit the most. Although the study focussed on the effects of solfège on music memorisation, the process of solfège lessons, taking place within the framework of the experiment, revealed the potential of solfège to benefit musicians, even the most advanced amongst them.

Moreover, the present research can constitute an important reference text for solfège teachers; apart from confirming their dedication to the specific method of educating musicians, it is always useful to apply a method being aware of the potential variety of its implications. Furthermore, the study showed that there were solfège subjects who

were not comfortable with solfège and were reluctant to use it in the tasks: this attitude indicates that, at least for some students, solfège was more a burden than it was a necessity; such findings are very likely a call for solfège teachers to alter their approach to teaching and exploit the vast variety of modern teaching methods suggested by modern pedagogy.

Certain useful conclusions from the experiments pertain to the work of instrumental teachers as well: although there is already a large body of literature on music memorisation, findings from the present study underlined the urgent need to link theory to practice. As many students appeared to be aware of the 'correct' answers regarding memorisation strategies, but subsequently failed to apply them in their learning of the pieces, it seems that instrumental and vocal teachers need to invest some time in putting memorisation theory into action and perhaps even devising new ways to fruitfully incorporate effective memorisation into everyday music practice.

Finally, despite their numerous methodological restrictions, the automated working memory experiments provided useful insights into the function of pitch labels within working memory, as well as a useful framework for further research in the field. Possible directions for future research could include further investigation of the nature of solfège syllables, letter names and pitch labels in general; comparison between pitch labels within visual and auditory working memory for music; and the examination of the level of integration between musical and verbal elements in the case of pitch labels.

REFERENCES

Acheson, D. J., Postle, B. R. and MacDonald, M. C. (2010) 'The interaction of concreteness and phonological similarity in verbal working memory', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36/1: 17–36.

Adler, M.S. (1975) 'Instrumentalists as Vocalists', *Music Educators Journal*, 61/7: 7–8.

Aiello, R. (1994) 'Music and Language: Parallels and Contrasts', in R. Aiello and J. Sloboda (eds), *Musical Perceptions* (pp. 40–63). Oxford: University Press.

Aiello, R. and Williamon, A. (2002) 'Memory', in R. Parncutt and G. McPherson (eds), *The science and psychology of music performance* (pp. 167–181). New York: Oxford University Press.

Andrade, J. and Donaldson, L. (2007) 'Evidence for an olfactory store in working memory', *Psychologia*, 50/2: 76–89.

Arnold, D. (1962-1963) 'Orphans and Ladies: The Venetian Conservatoires (1680–1790)', *Proceedings of the Royal Musical Association*, 89: 31–47.

Baddeley, A. D. and Hitch, G. J. (1974) 'Working memory', in G. H. Bower (ed), *The psychology of learning and motivation*, vol. 8 (pp. 47–90). New York: Academic Press.

Baddeley, A.D. (1990) *Human Memory: Theory and Practice*. Hove: Erlbaum.

Baddeley, A.D. (1999) *Essentials of Human Memory*. Hove: Psychology Press.

Baddeley, A.D. (2000) 'The Episodic Buffer: A New Component of Working Memory?', *Trends in Cognitive Sciences*, 4/11: 417–423.

Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J. and Freimer, N.B. (1998) 'Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components', *American Journal of Human Genetics*, 62: 224–231.

Bahrick, H. P. (1996) 'Synergistic strategies or memory research', in D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel and M.K. Johnson (eds), *Basic and Applied Memory Research: Theory in Context*, vol. 1 (pp. 51–62). New Jersey: Lawrence Erlbaum.

Bartlett, S.J. and Suber, P. (eds), (1987) *Self-Reference: Reflections on Reflexivity*. The Hague: Martinus Nijhoff.

Beardsley, M.C. (1971) 'On Aiken's Defense of the Humanities', *Journal of Aesthetic Education*, 5/4: 72–76.

Beattie, J.W. and Curtis, L.W. (1941) 'South American Music Pilgrimage: I. Colombia', *Music Educators Journal*, 28/2: 14–17.

Bentley, A. (1959) 'Fixed or Movable Do?', *Journal of Research in Music Education*, 7/2: 163–8.

Benfield, R. (1944) 'Musings', *Music Educators Journal*, 31/1: 60–61.

Berlin, B. and Kay, P. (1969) *Basic Color Terms: Their Universality and Evolution.* Berkeley: University of California Press.

Berz, W. (1995) 'Working Memory in Music: A Theoretical Model', *Music Perception*, 12/3: 353–364.

Bever, T.G., and Chiarello, R.J. (1974) 'Cerebral dominance in musicians and nonmusicians', *Science,* 185: 137–9.

Blacking, J. (1976) *How musical is man?* London: Faber and Faber.

Blasius, L.D. (1996) 'The mechanics of sensation and the construction of the Romantic musical experience', in I. Bent (ed.), *Music Theory in the Age of Romanticism* (pp. 3–24). Cambridge: Cambridge University Press.

Blethen, E. (1953) 'Music Education versus Solfeggio', *Music Educators Journal*, 39/4: 62–64.

Bloom, P. and Keil, F.C. (2001) 'Thinking through language'. *Mind &Language*, 16: 351–367.

Bonnel, A.-M., Faita, F., Peretz, I. and Besson, M. (2001) 'Divided Attention between Lyrics and Tunes of Operatic Songs: Evidence for Independent Processing', *Perception & Psychophysics*, 63/7: 1201–1213.

Bourassa, D.C. and Besner, D. (1994) 'Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists', *Psychonomic Bulletin & Review*, 1: 122–5.

Bourdeaux, M.J. (1975) *25 Leçons progressives de lecture de notes et de solfège rhythmique*. Paris: Gérard Billaudot.

Bower, G.H. (1998) 'An Associative theory of Implicit and Explicit Memory', in M.A. Conway, S.E. Gathercole and C. Cornoldi (eds), *Theories of Memory*, vol. 2 (pp. 25–60). Sussex: Psychology Press.

Brand, E. (2000) 'Children's mental musical organisation as highlighted by their singing errors', *Psychology of Music*, 28/1: 62–80.

Brandler, S. and Rammsayer, T.H. (2003) 'Differences in mental abilities between musicians and non-musicians', *Psychology of Music*, 31/2: 123–38.

Brown, K.D. (2003) 'An Alternative Approach to Developing Music Literacy Skills in a Transient Society', *Music Educators Journal*, 90/2: 46–54.

Bruce, V., Green, P.R. and Georgeson, M.A. (eds) (1996) *Visual Perception: Physiology, Psychology and Ecology.* Hove: Psychology Press.

Buckner, R.T. (1982) 'A History of Music Education in the Black Community of Kansas City, Kansas, 1905-1954', *Journal of Research in Music Education*, 30/2: 91–106.

Bullis, C. (1936) 'Harmony Study That Correlates with Musical Needs', *Music Educators Journal*, 23/3: 26–27.

Butler, D. and Brown, H. (1994) 'Describing the Mental Representation of Tonality in Music', in R. Aiello and J. Sloboda (eds), *Musical Perceptions* (pp. 191–212) Oxford: University Press.

Calvocoressi, M.D. (1932) 'Music in the Foreign Press', *The Musical Times*, 73/1073: 616–7.

Carmichael, L. C., Hogan, H. P. and Walters, A. A. (1932) 'An experimental study of the effect of language on the reproduction of visually perceived form' *Journal of Experimental Psychology*, 15: 73–86.

Carruthers, P. (2002) 'The cognitive functions of language', *Behavioral and Brain Sciences,* 25:657–674.

Castaldo, J. (1969) 'Creativity Can End Our Musical Isolationism', *Music Educators Journal*, 56/3: 36–38.

Chaffin, R. and Imreh, G. (1997) "Pulling teeth and torture': Musical memory and problem solving', *Thinking and Reasoning*, 3/4: 315–336.

Chaffin, R., Imreh, G. and Crawford, M. (2002) *Practicing perfection: memory and piano performance*. Mahwah, N.J.: Lawrence Erlbaum.

Chaffin, R., Imreh, G., Lemieux, A.F. and C. Chen (2003) '"Seeing the Big Picture": Piano Practice as Expert Problem Solving', *Music Perception*, 20/4: 465–490.

Chaffin, R. (2007) 'Learning *Claire de Lune*: Retrieval Practice and Expert Memorisation', *Music Perception*, 24/4: 377–393.

Chase, W.G. and Ericsson, K.A. (1982) 'Skill and working memory', in G.H. Bower (ed), *The psychology of learning and motivation*, vol. 16 (pp. 1–58). New York: Academic Press.

Chin, C.S. (2003) 'The development of absolute pitch: A theory concerning the roles of music training at an early developmental age and individual cognitive style', *Psychology of Music*, 31/2: 155–171.

Choksy, L. (1969) 'Kodály in and out of context', *Music Educators Journal*, 55/8: 57–59.

Chosky, L. (1981) *The Kodály context: creating an environment for musical learning*, Englewood Cliffs: Prentice-Hall.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1971) *Selected Readings* (J.P.B. Allen and P. van Bunen, eds). London: Oxford University Press.

Clark, A. (1989) *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: The MIT Press.

Cœuroy, A. and Baker, T. (1927) 'Present Tendencies of Sacred Music in France', *The Musical Quarterly*, 13/4: 582–604.

Coffman, D.D. (1987) 'Vocal Music and the Classroom Teacher, 1885 to 1905', *Journal of Research in Music Education*, 35/2: 92–102.

Conway, A.R.A, Kane, M.J., Bunting, M.F., Hambrick. Z.F., Wilhelm, O. and Engle, R. (2005) 'Working memory span tasks: A methodological review and user's guide', *Psychonomic Bulletin and Review*, 12/5: 769–786.

Craik, F.I.M. and Lockhart, R.S. (1972) 'Levels of Processing: A Framework for Memory Research', *Journal of Verbal Learning and Verbal Behavior*, 11/6: 671.

Craik, F.I.M. and Watkins, M.J. (1973) 'The Role of Rehearsal in Short-Term Memory', *Journal of Verbal Learning and Verbal Behavior*, 12/6: 599.

Craik, F.I.M. and Tulving, E. (1975) 'Depth of Processing and the Retention of Words in Episodic Memory', *Journal of Experimental Psychology: General*, 104/3: 268–294.

Craik, F.I.M., Naveh-Benjamin M. and Anderson D.N. (1996) 'Encoding and retrieval processes: similarities and differences', in M.A. Conway, S.E. Gathercole and C. Cornoldi (eds), *Theories of Memory*, vol. 2 (pp. 61–86). Sussex: Psychology Press.

Curran, T. and Schacter, D.L. (1996) 'Implicit Memory and Perceptual Brain mechanisms', in D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel and M.K. Johnson (eds), *Basic and Applied Memory Research: Theory in Context*, vol. 1 (pp. 221–240). New Jersey: Lawrence Erlbaum.

Curtiss, M. (1952) 'Gounod Before "Faust" ', *The Musical Quarterly*, 38/1: 48–67.

d' Arangues, M. (2004) 'A Special Homage: Alirio Diaz's Eightieth Birthday', *Latin American Music Review / Revista de Música Latinoamericana*, 25/2: 232–241.

Dannhäuser, A., Lavignac, A. and Lemoine, H. (1891) *Solfège des Solfeges*. Milwaukee: G. Schirmer.

Deliège, I. and Sloboda, J.A. (1997) *Perception and Cognition of Music.* Hove: Psychology Press.

Demorest, S.M., Morrison, S.J., Beken, M.N. and Jungbluth, D. (2008) 'Lost in translation: an enculturation effect in music memory performance', *Music Perception*, 25/3: 213–223.

Deutsch, D. (1999) 'The processing of pitch combinations', in D. Deutsch (ed), *The Psychology of Music*, 2nd edition (pp. 349–411). San Diego: Academic Press.

Deutsch, D. (ed) (1999) *The Psychology of Music*, 2nd edition. San Diego: Academic Press.

Dobszay, L. (1972) 'The Kodály method and its musical basis', *Studia Musicologica Academiae Scientiarum Hungaricae*, 1/4: 15–33.

Drake, C., McAdams, S. and Berthoz, A. (1999) 'Learning to sing a novel piece of music facilitates playing it on the violin but not the other way around: Evidence from performance segmentations', *Journal of the Acoustical Society of America*, 106/4: 2285.

Edlund, L. (1963) *Modus Novus*. Stockholm: Wilhelm Hansen.

Edwards, R.D. and Hodges, D.A. (2007) 'Neuromusical Research: An Overview of the Literature', in W. Gruhn and F.H. Rauscher (eds), *Neurosciences in Music Pedagogy* (pp. 1–25). New York: Nova Science Publishers.

Elliott, C.A. (1982) 'The Music-Reading Dilemma', *Music Educators Journal*, 68/6: 33–34+59–60.

Engle, R.W. (2001) 'What is working-memory capacity?', in H.L. Roediger and J.S. Nairne (eds), *The Nature of Remembering: Essays in Honor of Robert G. Crowder* (pp. 297–314). Washington, DC: American Psychological Association.

Erceg-Hurn, D.M. and Mirosevich, V.M. (2008) 'Modern robust statistical methods: An easy way to maximize the accuracy and power of your research', *American Psychologist*, 63/7: 591–601.

Ericsson, K.A. (1985) 'Memory Skill', *Canadian Journal of Psychology*, 39/2: 188–231.

Ericcson, K.A. and Kintsch, W. (1995) 'Long-Term Working Memory', *Psychological Review*, 102/2: 211–245.

Eysenck, M.W. (1977) *Human Memory: Theory, Research and Individual Differences.* Oxford: Pergamon.

Eysenck, M.W. and Keane, M.K. (2000) *Cognitive Psychology: A student's handbook*, 4th edition. Hove and New York: Psychology Press.

Field, A. (2009) *Discovering statistics using SPSS*, 3rd edition. London: SAGE.

Fine, P., Berry, A. and Rosner, B. (2006) 'The effect of pattern recognition and tonal predictability on sight-singing ability', *Psychology of Music*, 34/4: 431–447.

Fodor, J.A. (1975) *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J.A. (1983) *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J.A. (2008) *LOT 2: The Language of Thought Revisited*. Oxford: Clarendon Press.

Forssmark, K.A. (1941) 'A Method for Teaching the Reading of Rhythms', *Music Educators Journal*, 27/6: 60–62.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R. and Pantev C. (2005) 'Automatic Encoding of Polyphonic Melodies in Musicians and Nonmusicians', *Journal of Cognitive Neuroscience*, 17/10: 1578–1592.

Gardner, H. (1983) *Frames of mind: The Theory of Multiple Intelligences.* London: Heinemann.

Gathercole, S. E., Pickering, S. J., Hall, M. and Peaker, S. M. (2001) 'Dissociable lexical and phonological influences on serial recognition and serial recall', *The Quarterly Journal of Experimental Psychology*, 54A/1: 1–30.

Gehrkens, K.W. (1960) 'The Development of a College Curriculum in Music Education', *Music Educators Journal*, 47/2: 31–34.

Gerson-Kiwi, E. (1958) 'Musicology in Israel', *Acta Musicologica*, 30/1: 17–26.

Gilman, E., Underwood, G. and Morehen, J. (2002) 'Recognition of visually presented musical intervals', *Psychology of Music*, 30: 48–57.

Ginsborg, J. (2002) 'Classical singers learning and memorising a new song: An observational study', *Psychology of music*, 30/1: 58–101.

Ginsborg, J. (2004) 'Strategies for memorizing music', in A. Williamon (ed), *Musical excellence: Strategies and techniques to enhance performance* (pp. 123–142). Oxford: Oxford University Press.

Ginsborg, J. and Sloboda, J.A. (2007) 'Singers' recall for the words and melody of a new, unaccompanied song', *Psychology of Music*, 35/3: 421–440.

Gleitman, L. and Papafragou, A. (2005) 'Language and thought', in K. Holyoak and B. Morrison (eds), *Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge, England: Cambridge University Press.

Goldstone, R.L. (1994) 'The role of similarity in categorization: providing a groundwork', *Cognition*, 52/2: 125–157.

Goldstone, R.L. and Barsalou, L.W. (1998) 'Reuniting perception and conception', *Cognition*, 65: 231–262.

Gordon, I.E. (1997) *Theories of Visual Perception.* Chichester: Wiley.

Graham, G. (1997) *Philosophy of the Arts: An introduction to Aesthetics*. London: Routledge.

Gregersen, P.K., Kowalsky, E., Kohn, N. and Marvin, E.W. (1999) 'Absolute Pitch: Prevalence, Ethnic Variation, and Estimation of the Genetic Component', *American Journal of Human Genetics*, 65: 911–3.

Gregersen, P.K., Kowalsky, E., Kohn, N. and Marvin, E.W. (2000) 'Early Childhood Music Education and Predisposition to Absolute Pitch: Teasing Apart Genes and Environment', *American Journal of Medical Genetics*, 98: 280–2.

Gruhn, W. and Rauscher, F.H. (eds) (2007) *Neurosciences in Music Pedagogy.* New York: Nova Science Publishers.

Gruneberg, M.M., Morris, P.E., Sykes, R.N. and Herrmann, D.H. (1996) 'The practical application of memory research: practical problems in the relationship between theory and practice', in D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel and M.K. Johnson (eds), *Basic and Applied Memory Research: Theory in Context*, vol. 1 (pp. 63–82). New Jersey: Lawrence Erlbaum.

Hallam, S. (1997) 'The development of memorisation strategies in musicians: implications for education', *British Journal of Music Education*, 14/1: 87–97.

Hallam, S. (2005) *Music psychology in education*. London: Institute of Education Press.

Halpern, A.R., Bartlett, J.C. and Dowling, W.J. (1995) 'Aging and experience in the recognition of musical transposition', *Psychology and Aging*, 10: 325–342.

Hardin, C.L. and Maffi, L. (eds) (1997) *Color Categories in Thought and Language*. Cambridge: Cambridge University Press.

Hartley, T. and Houghton, G. (1994) 'A linguistically-constrained model of short-term memory for nonwords', *Journal of Memory and Language*, 35: 1–31.

Heaton, P. (2003) 'Pitch memory, labelling and disembedding in autism', *Journal of Child Psychology and Psychiatry*, 44/4: 543–551.

Hébert, S., Racette, A., Gagnon, L. and Peretz, I. (2003) 'Revisiting the Dissociation between Singing and Speaking in Expressive Aphasia', *Brain*, 126: 1838–1850.

Henderson, M.T., Crews, A. and Barlow, J. (1945) 'A study of the effect of music distraction on reading efficiency', *Journal of Applied Psychology*, *29*:313–7.

Hiatt, J.S. and Cross, S. (2006) 'Teaching and Using Audiation in Classroom Instruction and Applied Lessons with Advanced Students', *Music Educators Journal*, 92/5: 46–49.

Hirst, W. and Pinner, E. (1996) 'Memory and Attention', in D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel and M.K. Johnson (eds), *Basic and Applied Memory Research: Theory in Context*, vol. 1 (pp. 241–258). New Jersey: Lawrence Erlbaum.

Hood, M. (1960) 'The challenge of "Bi-Musicality"', *Ethnomusicology*, 4/2: 55–59.

Hoskins, W. (1959) 'Re: The Eastman "Sight-Singing Manual"', *Journal of Music Theory*, 3/1: 150–154.

Houlahan, M. and Tacka, P. (2008) *Kodály today: A cognitive approach to elementary music education*. New York: Oxford University Press.

Hughes, A. and Gerson-Kiwi, E.: 'Solmization.' *Grove Music Online. Oxford Music Online*, available at: http://www.oxfordmusiconline.com/subscriber/article/grove/music/26154, [accessed 23/03/2006]

Hughes, D.W. (2000) 'No nonsense: The logic and power of acoustic-iconic mnemonic systems', *British Journal of Ethnomusicology*, 9/2: 93–120.

Hughes, E. (1915) 'Musical memory in piano playing and piano study', *Musical Quarterly*, 1: 592–603.

Hulme, C., Maughan, S. and Brown, G. D. A. (1991) 'Memory for familiar and unfamiliar words: Evidence for a longer term memory contribution to short-term memory span', *Journal of Memory and Language*, 30: 685–701.

Humphreys, J.T. (2006) 'Senior Researcher Award Acceptance Address: Observations about Music Education', *Journal of Research in Music Education*, 54/3: 183–202.

Hutton, P.H. (1987) 'The Art of Memory Reconceived: From Rhetoric to Psyhoanalysis', *Journal of the History of Ideas*, 48/3: 371–392.

Imberty, M. (1993) 'How Do We Perceive Atonal Music? Suggestions For a Theoretical Approach', *Contemporary Music Review*, 9/1&2: 325–337.

Imberty, M. (1996) 'Linguistic and musical development in preschool and school-age children', in I. Deliège and J.A. Sloboda (eds.), *Musical beginnings: Origins and development of musical competence*, (pp. 191–213). New York: Oxford University Press.

Inglefield, B. (1962) 'Music in the Public Schools of Puerto Rico', *Music Educators Journal*, 48/5: 86+88.

International Phonetic Association (1999) *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Janata, P. and Grafton, S.T. (2003) 'Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music', *Nature Neuroscience*, 6: 682–7.

Jean-Aubry G. (1918) 'Claude Debussy' (Trans. F.H. Martens), *The Musical Quarterly*, 4/4: 542–554.

Johnson, A.J. and Miles, C. (2009) 'Single-probe serial position recall: Evidence of modularity for olfactory, visual, and auditory short-term memory', *The Quarterly Journal of Experimental Psychology*, 62/2: 267–275.

Jones, A. (1966) 'The Tone-Word System of Carl Eitz', *Journal of Research in Music Education*, 14/2: 84–98.

Kalakoski, V. (2007) 'Effect of skill level on recall of visually presented patterns of musical notes', *Scandinavian Journal of Psychology*, 48: 87–96.

Kendel, J.C. (1955) 'Keyboard experience: A foundation for musical knowledge and enjoyment', *Music Educators Journal*, 41/5: 38+42.

Knecht, M. G. (2003) 'Music Expertise and Memory: the relationship between music expertise and memory of music patterns, within various degrees of contextual constraint', *Music Education Research*, 5/3: 227–242.

Kodály, Z. (1941) *Tizenöt Kétszólamú Énekgyakorlat*. Budapest: Editio Musica.

Koelsch S., Gunter, T.C., Cramon, Y., Zysset, S., Lohmann, G. and Friederici, A.D. (2002) 'Bach Speaks: A Cortical "Language-Network" Serves the Processing of Music', *NeuroImage*, 17/2: 956–966.

Kolers, P.A. (1973) 'Remembering Operations', *Memory & Cognition*, 1/3: 347–355.

Kotts-Murphy, M. and Brown, T.S. (1986) 'A Comparison of Preferences for Instructional Objectives between Teachers and Students', *Journal of Research in Music Education*, 34/2: 134–9.

Krampe, R. and Ericsson, K. (1995) 'Deliberate practice and elite musical performance', in J. Rink (ed), *The practice of performance: Studies in musical interpretation* (pp. 84–104). Cambridge: Cambridge University Press.

Kramer, L. (2002) 'Signs Taken for Wonders: Words, Music, and Performativity', *Word and Music Studies*, 4: 35–47, available at: http://www.ingentaconnect.com/content/rodopi/wms/2002/00000004/00000001/art00004, [accessed 12/01/06]

Krumhansl, C.L. (1990) *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.

Ladukin, N. (1980) *Monophonic solfège*. Moscow: Editio Musica.

Lamers, W.M. (1960) 'The Two Kinds of Music', *Music Educators Journal*, 46/4: 36–38+82.

Lawler, V. (1947) 'Music Education Reaches 'round the World', *Music Educators Journal*, 33/6: 18–20.

Leonhard, C. (1982) 'Humanizing Music in a Mechanized Society', *Music Educators Journal*, 68/9: 23–24.

Levitin, D.J. (1994) 'Absolute memory for absolute pitch: Evidence from the production of learned melodies', *Perception & Psychophysics*, 56/4: 414–423.

Levitin, D.J. and Zatorre, R.J. (2003) 'On the Nature of Early Music Training and Absolute Pitch: A Reply to Brown, Sachs, Cammuso, and Folstein', *Music Perception*, 21/1: 105–110.

Levitin, D.J. and Menon, V. (2003) 'Musical Structure is Processed in 'Language' Areas of the Brain: A Possible Role for Brodmann Area 47 in Temporal Coherence', *NeuroImage*, 20/4: 2142–2152.

Levitin, D.J. and Rogers, S.E. (2005) 'Absolute Pitch: Perception, Coding and Controversies', *Trends in Cognitive Sciences*, 9/1: 26–33.

Lidji, P., Peretz, I., Kolinsky, R., Colin, C. and Morais, J. (2007, May) *Pre-attentive and attentive change detection of vowel and pitch in sung syllables: An event-related potential study.* Paper presented at Language and Music as Cognitive Systems, Cambridge (UK).

Lim, S. and Lippman, L.G. (1991) 'Mental Practice and Memorisation of Piano Music', *The Journal of General Psychology*, 118/1: 21–30.

Long, J.H. (1950) 'Shakespeare and Thomas Morley', *Modern Language Notes*, 65/1: 17-22.

Long, P.A. (1977) 'Relationships Between Pitch Memory in Short Melodies and Selected Factors', *Journal of Research in Music Education*, 25/4: 272–282.

Lowe, A. and Pryor, H.S. (1959) 'Music Education in the Union of Soviet Socialist Republics', *Music Educators Journal*, 45/6: 28–32.

Lowens, I. and Lombard, L. (1985) 'Louis Lombard's "Our Conservatories" (1892)', *American Music*, 3/3: 347–351.

Lupyan, G. (2008a) 'The conceptual grouping effect: Categories matter (and named categories matter more)', *Cognition*, 108: 566–577.

Lupyan, G. (2008b) 'From Chair to "Chair": A Representational Shift Account of Object Labeling Effects on Memory', *Journal of Experimental Psychology: General*, 137/2: 348–369.

Lutgen, B. (1889) 'At What Age May Girls Begin to Learn Singing?', *The Musical Times and Singing Class Circular*, 30/560: 620.

Madsen, C.K. and Madsen, K. (2002) 'Perception and cognition in music: Musically trained and untrained adults compared to sixth-grade and eighth-grade children', *Journal of Research in Music Education*, 50/2: 111–130.

Magne, C., Aramaki, M., Astesano, C., Gordon, R.L., Ystad, S., Farner, S., Kronland-Martinet, R. and Besson, M. (2005) 'Comparison of Rhytmic Processing in Language and Music- An Interdisciplinary Approach', *The Journal of Music and Meaning*, [Online], 3/5, available at: http://www.musicandmeaning.net/issues/showArticle.php?artID=3.5 [accessed 23/12/2005]

Mangino, B. (1963) 'Why Italians Love Good Music', *Music Educators Journal*, 49/3: 47–48.

Mann, R. (1989) 'Why Should Elementary Students Have All the Fun?', *Music Educators Journal*, 76/1: 39–42.

Marin, O.S.M. (1989) 'Neuropsychology, Mental Cognitive Models and Music Processing', *Contemporary Music Review*, 4: 255–263.

Marin, O.S.M. and Perry, D.W. (1999) 'Neurological Aspects of Music Perception and Performance', in D. Deutsch (ed), *The Psychology of Music*, 2nd edition (pp. 653–724). San Diego: Academic Press.

Marraro, H.R. (1946) 'Italian Music and Actors in America During the Eighteenth Century', *Italica*, 23/2: 103–117.

Martin, B.A. (1991) 'Effects of Hand Signs, Syllables, and Letters on First Graders' Acquisition of Tonal Skills', *Journal of Research in Music Education*, 39/2: 61–170.

Martin, R.C., Wogalter, M.S. and Forlano, J.G. (1988) 'Reading Comprehension in the Presence of Unattended Speech and Music', *Journal of Memory and Language*, 27/4: 382–398.

Matthay, T. (1926) *On memorizing and playing from memory and on the laws of practice generally*. London: Oxford University Press.

McClelland, J.L. and Lupyan, G. (2002) 'Double dissociations never license simple inferences about underlying brain organization, especially in developmental cases',

Commentary on Thomas and Karmiloff-Smith, Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling, *Behavioral and Brain Sciences*, 25/6: 763–4.

McNaught, W. (1931) 'About Tonic Sol-Fa', *The Musical Times*, 72/1065: 977–982.

McNaught, W. G. (1892) 'The History and uses of Sol-Fa Syllables', *Journal of the Royal Musical Association*, 19/1: 35–51.

Merriam, A.P. (1964) *The Anthropology of Music.* Chicago: Northwestern University Press.

Miller, S.D. (1973) 'Guido d'Arezzo: Medieval Musician and Educator', *Journal of Research in Music Education*, 21/3: 239–245.

Miranda, M.L. (2000) 'Developmentally Appropriate Practice in a Yamaha Music School', *Journal of Research in Music Education*, 48/4: 294–309.

Mishkin, M. and Appenzeller, T. (1987) 'The Anatomy of Memory', *Scientific American*, 256: 80–89.

Mishra, J. and Backlin, W. (2007) 'The effects of altering environmental and instrumental context on the performance of memorized music', *Psychology of Music*, 35/3: 453–472.

Mithen, S. (2005) *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld and Nicolson.

Morphy, E.W. (1918) 'Violin teaching in its relation to the organization of civic orchestras', *The Musical Quarterly*, 4/1: 50–60.

Moscovitch, M., Kapur, S., Kohler, S and Houle, S. (1995) 'Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography study in humans (visual cortex/hippocampus)', *Proceedings of the National Academy of Sciences USA: Psychology*, 92: 3721–5.

Münte, T.F., Altenmüller, E. and Jäncke , L. (2002) 'The musician's brain as a model of neuroplasticity', *Nature Reviews Neuroscience*, 3: 473–8.

Murdock, B.B. Jr. (1967) 'Recent Developments in Short-Term Memory', *British Journal of Psychology*, 58: 421–433.

Neisser, U. (1976) *Cognition and Reality*. San Francisco: Freeman and Company.

Nettl, B.: 'Music.', *Grove Music Online. Oxford Music Online*, available at: http://www.oxfordmusiconline.com/subscriber/article/grove/music/40476, [accessed 12/6/2009]

Newton, M. (1953) 'Music for the Elementary Teacher', *Music Educators Journal*, 40/1: 67–68.

Nomme, A. (1954) 'The Bosses Need a Boss', *Music Educators Journal*, 40/4: 58+61.

Nye, R.E. (1948) 'Syllables, Numbers and Sight Reading', *Music Educators Journal*, 35/2: 52–53.

Ohanian, J. (1977) Review of *The Tender Tyrant: Nadia Boulanger, A Life Devoted to Music*, by Alan Kendall, *Notes, Second Series*, 34/2: 354–6.

Osmond-Smith, D. (1989) 'Between Music and Language: A View from the Bridge', *Contemporary Music Review*, 4: 89–100.

O'Neill, D.D. (1948) 'Harmony in High School', *Music Educators Journal*, 35/2: 62–63.

Oura, Y. and Hatano, G. (2004) 'Parsing and memorizing tonal and modal melodies', *Japanese Psychological Research*, 46/4: 308–321.

Palmer, C. (2005) 'Sequence Memory in Music Performance', *Current Directions in Psychological Science*, 14/5: 247–250.

Papke, R.E. (1970) 'The New Breed of Band Director Thinks Comprehensively', Music *Educators Journal*, 57/3: 40–42+105+107–8

Peithman, H.W. (1945) 'Student Aids to Student Achievements', *Music Educators Journal*, 31/4: 32+37+61–63.

Peles, S. (1994) 'A conversation with Claudio Spies: Interview with Claudio Spies', *Perspectives on New Music*, 32/1: 292–325.

Peretz, I. and Morais, J. (1989) 'Music and Modularity', *Contemporary Music Review*, 4: 279–293.

Peretz, I. and Coltheart, M. (2003) 'Modularity of music processing', *Nature Neuroscience*, 6: 688–691.

Peretz, I. and Hyde, K. L. (2003) 'What is Specific to Music Processing? Insights form Congenital Amusia', *Trends in Cognitive Sciences*, 7/8: 362–7.

Peretz, I. and Zatorre, R.J. (eds) (2003) *The cognitive neuroscience of music.* Oxford: University Press.

Peretz, I., Gagnon, L., Hébert, S. and Macoir, J. (2004) 'Singing in the Brain: Insights from Cognitive Neuropsychology', *Music Perception*, 21/3: 373–390.

Peretz, I., Radeau, M. and Arguin, M. (2004) 'Two-way Interactions between Music and Language: Evidence from Priming Recognition of Tune and Lyrics in Familiar Songs', *Memory & Cognition*, 32/1: 142–152.

Peretz, I. (2007, May) *Comparisons between Music and Language as Tests of Modularity.* Paper presented at Language and Music as Cognitive Systems, Cambridge (UK).

Perkins, D.N. (1986) *Knowledge as design.* London: Lawrence Erlbaum.

Peters, G.D. and Miller, R.F. (1982) *Music Teaching and Learning.* New York: Longman.

Perry, D.W. (1993) 'A cognitive Neuropsychological Analysis of Melody Recall', *Contemporary Music Review*, 9/1&2: 97–111.

Phillips, P. (1997) 'Review: Venice Preserved', *The Musical Times*, 138/1784: 32–33.

Phillips, I. and Martens, F.H. (1920) 'The French National Conservatory of Music', *The Musical Quarterly*, 6/2: 214–226.

Poirier, M. and Saint-Aubin, J. (1995) 'Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall', *Quarterly Journal of Experimental Psychology*, 48A: 384–404.

Prod'homme, J.G. (1918) 'Claude Achille Debussy' (Trans. M. Barton), *The Musical Quarterly*, 4/4: 555–571.

Profita, J. and Bidder, T.G. (1998) 'Perfect Pitch', *American Journal of Medical Genetics*, 29/4: 763–771.

Proudfit, M.E. (1966) 'Music Education in Central America', *Music Educators Journal*, 52/6: 110–111.

Radocy, R.E. and Boyle, J.D. (1979) *Psychological Foundations of Musical Behavior.* Illinois: Charles C. Thomas.

Rappaport, J.C. (2004) 'Exception Taken', *Music Educators Journal*, 90/4: 9–11.

Repovs, G. and Baddeley, A.D. (2006) 'The multi-component model of working memory: explorations in experimental cognitive psychology', *Neuroscience*, 139/1: 5–21.

Review of *Elementary Course for Vocalising and Pronouncing the English Language*, by N. Macfarren. (1868) *The Musical Times and Singing Class Circular*, 13/301: 323.

Revelli, W.D. (1961) 'Reflections on a Musical Adventure. Part I: The Soviet Union', *Music Educators Journal*, 48/1: 37–41.

Richter, A. (1938) 'Music and the forgotten pupil', *Music Educators Journal*, 25/2: 19–22.

Rink, J. (ed) (2002) *Musical Performance: A Guide to Understanding*. Cambridge: Cambridge University Press.

Rogers, L., Hallam, S., Creech, A. and Preti, C. (2008) 'Learning about what constitutes effective training from a pilot programme to improve music education in primary schools', *Music Education Research*, 10/4: 485–497.

Rolland, P. (1947) 'The Teaching of Strings', *Music Educators Journal*, 33/5: 34+36+38+58–59.

Roth, I. (1986) 'Conceptual Categories', in I. Roth and Frisby, J. (eds) *Perception and Representation: a Cognitive Approach* (pp. 23–58). Milton Keynes: Open University Press.

Roth, I. and Frisby, J.P. (eds) (1986) *Perception and Representation: a Cognitive Approach*. Milton Keynes: Open University Press.

Rubin, E. (1990) 'Jeannette Meyers Thurber and the National Conservatory of Music', American Music, 8/3: 294–325.

Ruthsatz, J. and Detterman, D.K. (2003) 'An extraordinary memory: The case study of a musical prodigy', *Intelligence*, 31: 509–518.

Salamé, P. and Baddeley, A. (1989) 'Effects of background music on phonological short-term memory', *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 41/1: 107–122.

Samson, S. and Zatorre, R.J. (1991) 'Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17/4: 793–804.

Sands, M. (1942) 'The Singing-Master in Eighteenth-Century England', *Music & Letters*, 23/1: 69–80.

Sands, M. (1943-1944) 'The Teaching of Singing in Eighteenth Century England', *Proceedings of the Musical Association*, 70: 11–33.

Saussure, F. de (1983) *Course in General linguistics* (Trans. R. Harris). London: Duckworth.

Schaaf, E. (1985) 'From Idea to Tradition: The Peabody Prep', *Music Educators Journal*, 72/1: 38–43.

Schwadron, A.A. (1971) 'On Words and Music: Toward an Aesthetic Conciliation', *Journal of Aesthetic Education*, 5/3: 91–108.

Seeger, C. (1944) 'Music Education in the Americas', *Music Educators Journal*, 30/6: 14–55.

Segalowitz, N., Cohen, P., Chan, A. and Prieur, T. (2001) 'Musical recall memory: Contribution of elaboration and depth of processing', *Psychology of Music*, 29/2: 139–148.

Seitz, J.A. (2005) 'Dalcroze, the body, movement and musicality', *Psychology of Music*, 33/4: 419–435.

Selfridge-Field, E. (1976) 'Bassano and the Orchestra of St Mark's', *Early Music*, 4/2: 153–8.

Shallice, T., Fletcher, P. and Dolan, R. (1998) 'The Functional Imaging of Recall', in M.A. Conway, S.E. Gathercole and C. Cornoldi (eds), *Theories of Memory*, vol. 2 (pp. 247–258). Sussex: Psychology Press.

Shamrock, M. (1997) 'Orff-Schulwerk: An integrated foundation', *Music Educators Journal*, 83/6: 41–44.

Sigmund, C. (1972) 'Segundo Luis Moreno (1882-1972): Ecuador's Pioneer Musicologist', *Anuario Interamericano de Investigacion Musical*, 8: 71–104.

Siegel, J.A. (1974) 'Sensory and verbal coding strategies in subjects with absolute pitch', *Journal of Experimental Psychology*, 103: 37–44.

Siler, H. (1956) 'Toward an International Solfeggio', *Journal of Research in Music Education*, 4/1: 40–43.

Sloboda, J.A. (1984) 'Experimental Studies of Music Reading: A Review', *Music Perception* 2/2: 222–36.

Sloboda, J.A. (1985) *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Oxford University Press.

Sloboda, J.A. (ed), (2000) *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*. Oxford: Clarendon Press.

Sloboda, J.A. (2005) *Exploring the Musical Mind: Cognition, Emotion, Ability, Function*. Oxford: Oxford University Press.

Sloboda, J.A., Wise, K.J. and Peretz, I. (2005) 'Quantifying tone deafness in the general population', *Annals of the New York Academy of Sciences*, 1060: 255–261.

Smith, M. (1934) 'Solfège: An Essential in Musicianship', *Music Supervisors' Journal*, 20/5: 16–61.

Smucker, M.D., Allan, J. and Carterette, B. (2007, November) *A comparison of statistical significance tests for information retrieval evaluation*. Paper presented at 16th ACM Conference on Information and Knowledge Management, Lisboa (Portugal).

Snowling, M., Chiat, S. and Hulme, C. (1991) 'Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley', *Applied Psycholinguistics*, 12: 369–373.
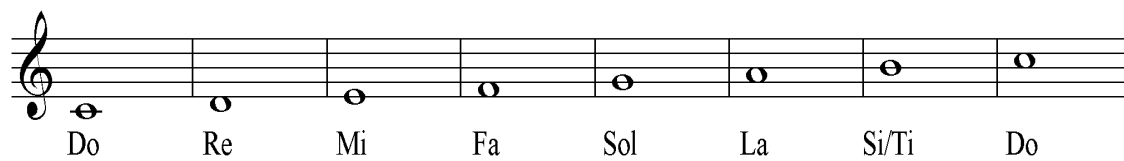
Squire, L. R., Shimamura, A. P. and Graf, P. (1985) 'Independence of recognition memory and priming effects: A neuropsychological analysis', *Journal of Experimental Psychology: Learning, Memory and Cognitio*n, 11: 37–44.

Stewart, L. and Williamon, A. (2008) 'What are the implications of neuroscience for music education?', *Educational Research*, 50/2: 177–186.

Swanwick, K. (1988) *Music, Mind, and Education*. London: Routledge.

Swanwick, K. (1999a) 'Music education: Closed or open?', *Journal of Aesthetic Education*, 33/4: 127–141.

Swanwick, K. (1999b) *Teaching Music Musically*. London: Routledge.

Takeuchi, A.H. and Hulse, S.H. (1993) 'Absolute Pitch', *Psychological Bulletin*, 113/2: 345–361.

Trehub, S. (2003) 'The developmental origins of musicality', *Nature Neuroscience*, 6: 669–673.

Troth, E.W. (1961) 'The Teacher-Training Program in Music at Chautauqua Institution, 1905-1930', *Journal of Research in Music Education*, 9/1: 37–46.

Tulving, E. and Thompson, D.M. (1973) 'Encoding Specificity and Retrieval Processes in Episodic Memory', *Psychological Review*, 80/5: 352–373.

Unsworth, N., Heitz, R.P., Schrock, J.C. and Engle, R.W. (2005) 'An automated version of the operation span task', *Behavior Research Methods*, 37/3: 498–505.

von Ende, R.C. (1965) 'Music and Shakespeare', *Music Educators Journal*, 51/3: 48–50.

Wade, N.J. and Swanston, M.T. (1991) *Visual Perception: An introduction.* London: Routledge.

Ward, W.D. (1999) 'Absolute Pitch', in D. Deutsch (ed), *The Psychology of Music*, 2nd edition (pp. 265–298). San Diego: Academic Press.

Wassel, A.W. (1966) 'The Bodra Smyana Choir: Interview with Bontcho Bontchev', *Music Educators Journal*, 52/5: 52–55+147–151.

Wayne, M. (1945) 'Music Education in Wartime France', *Music Educators Journal*, 31/6: 60–76.

Whorf, B.L. (1956) *Language, Thought, and Reality.* Ed. J.B.Carroll. Cambridge: M.I.T. Press.

Wilding, J. and Valentine, E. (1996) 'Memory expertise', in D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel and M.K. Johnson (eds), *Basic and Applied Memory Research: Theory in Context*, vol. 1 (pp. 300–320). New Jersey: Lawrence Erlbaum.

Williamon, A. (1999) 'The value of performing from memory', *Psychology of music*, 27/1: 84–95.

Williamon, A. (2002) 'Memorizing music', in J. Rink (ed), *Musical Performance* (pp. 113–126). Cambridge: Cambridge University Press.

Williamon, A. and Valentine, E. (2002) 'The Role of Retrieval Structures in Memorizing Music', *Cognitive Psychology*, 44: 1–32.

Williamon, A., Valentine, E. and Valentine, J. (2002) 'Shifting the focus of attention between levels of musical structure', *European Journal of Cognitive Psychology*, 14/4: 493–520.

Williams, D. B. (1975) 'Short-Term Retention of Pitch Sequence', *Journal of Research in Music Education*, 23/1: 53–66.

Williamson, V. J., Mitchell, T., Hitch, G. J. and Baddeley, A. D. (2010) 'Musicians' memory for verbal and tonal material under conditions of irrelevant sound', *Psychology of Music*, 38: 331–350.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D. and Stewart, L. (2010) 'Faster decline of pitch memory over time in congenital amusia', *Advances in Cognitive Psychology*, 6: 15–22.

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

Wilson, M.E. (1946) 'An Obsolescent Degree', *Journal of Higher Education*, 17/7: 344–394.

Wolbers, M. (2002) 'Singing in the band rehearsal', *Music Educators Journal*, 89/2: 37–41.

Wöllner, C., Halfpenny, E., Ho, S. and Kurosawa, K. (2003) 'The effects of distracted inner hearing on sight-reading', *Psychology of Music*, 31/4: 377–389.

Yantis, S. (ed) (2001) *Visual Perception: Essential Readings.* Hove: Psychology Press.

Zatorre, R.J. and Beckett, C. (1989), 'Multiple coding strategies in the retention of musical tones by possessors of absolute pitch', *Memory and Cognition*, 17/5: 582–589.

Zatorre, R.J., Perry, D., Becket, C., Westbury, C. and Evans, A.C. (1998) 'Functional anatomy of musical processing in listeners with absolute and relative pitch', *Proceedings of the National Academy of Sciences, 95*: 3172–7.

Παυλίδου, Θ. (1997) *Επίπεδα Γλωσσικής Ανάλυσης*. Θεσσαλονίκη: Παρατηρητής.
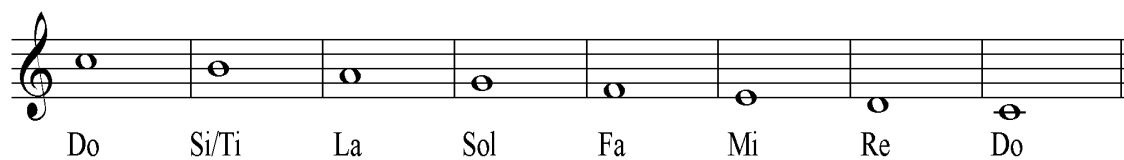
Appendix 1

Solfège lessons

Solfège handbook (Main study)

Do   Re   Mi   Fa   Sol   La   Si/Ti   Do

Do   Si/Ti   La   Sol   Fa   Mi   Re   Do

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

LAD (74)

11.

LAD (75)

12.

H.L. (I-88)

13.

14.

Andante

15.

Fine

16.

17.

18.

ROD. (I-103)*

19.

ROD. (I-102)*

20. Moderato. (♩ = 80)

mf

mf *p* *mf*

G.C. (I-93)

21.

H.L. (I-109)

22. Moderato (♩ = 104)

f

Fine *f*

mf *p*

poco a poco cresc.

ROD. (I-142)

23.

ROD. (I-115)

Moderato (♩ = 80)

24.

H.L. (I-145)

25.

Fine

cresc.

26.

27.

H.L. (I-108)*

28.

ROD. (I-57)*

29.

B (57)

30.

LEO. (I-116)

31.

G.C. (I-72)

32.

B (54)*

33.

37.

38.

39.

40.

41.

42.

43.

ROD (I-127)*

44.

E. H. Grieg

45. Allegretto

F. P. Schubert

46.

Exercises in Bass Clef

47.

H.L. (I-148)

Fine

48.

H.L. (I-149)

Fine

49.

H.L. (I-152)

50.

G.C. (I-158)

B (115)

51.

ROD. (I-166)

52.

H.L. (I-165)

53.

Fine

Exercises with changes between Treble and Bass Clefs

ROD (3A-13)

57.

SCHNEI (3A-15)

58.

SCHNEI (3A-18)

59.

60.

Moderato (♩ = 104)

H.L. (I-109)*

61.

H.L. (3A-16)

62.

Exercises in Alto Clef

Do   Re   Mi   Fa   Sol   La   Si/Ti   Do

Do   Si/Ti   La   Sol   Fa   Mi   Re   Do

ROD (A3-43)

63.

ROD. (I-103)*

64.

ROD (3A-45)

65.

69.

Rallent.

Exercises with changes between Treble, Bass and Alto Clefs

70.

71.

72.

ROD (3A-17)*

73.

ROD. (I-133)*

74.

H.L. (1-135)*

75.

Two-voice Exercises

A I, 964

76.

A I, 970

77.

A I, 981

78.

P V, 7

79.

K. Z. (1)

80.

Z. K. (6)

81.

P VI, 20

82.

83.

84.

Z. K. (8)

P VI

G. F. Händel

87.

Moderato

88.

89.

F. Durante

90.

Z. K. (3)

91.

92.

93.

94.

P VI, 30

95.

P VI, 21

96.

L III, 17

97.

98.

D III, 25

99.

P VI, 7

100.

P VI, 30

101.

Solfège lessons questionnaire

1. **Name:** _____

2. **Instrument (if more than one, please write them in order of years of experience, starting with the one that you are more experienced on):**

    _____

    _____

    3.**How would you rate the level of difficulty of the solfège lessons overall?**

    a.Very difficult
    b.Difficult
    c.Neither difficult nor easy
    d.Easy
    e.Very Easy

4. **Do you feel that you know solfège well enough to use it when reading/ practising an unfamiliar piece of music?**

    Yes   No

5. **If yes, would you use solfège when reading/practicing an unfamiliar piece of music?**

    Yes   No

6. **Did solfège improve or help any aspect of your personal music practice/ performance**?

    Yes   No

7. **If yes, what aspect was that and in what way did solfège help**?

    _____

    _____

    _____

    _____

8. How would you feel about solfège instruction being incorporated in a University curriculum?
      a.Approve strongly
      b.Approve
      c.Neither approve nor disapprove
      d.Disapprove
      e.Disapprove strongly

9. How would you feel about solfège instruction being incorporated in a school curriculum?
      a.Approve strongly
      b.Approve
      c.Neither approve nor disapprove
      d.Disapprove
      e.Disapprove strongly

10. Do you have any further comments regarding solfège and/ or solfège lessons?

Thank you for your time

Solfège lessons feedback form

Please circle the appropriate answer or write it in the space provided

1. **How would you rate your solfège skills at this level?**

   a. Extremely good
   b. Good
   c. Neither good nor poor
   d. Poor
   e. Extremely poor

2. **Did you feel you had enough time to learn solfège efficiently?**   Yes   No

3. **If no, could you give an estimation about how much more time you think you would require in order to learn solfège more satisfactorily?**

   a.2-4 weeks
   b.1 semester
   c.1 year
   d.more than 1 year

4. **Did you feel that you have learnt the material covered in the lessons sufficiently?**
                        Yes   No

5. **If no, why do you think this happened?**

   a.Due to the teacher

   b.Due to lack of individual practice

   c.Due to other students' response to the lessons

   d.Other (please specify)

   _____

   _____

   _____

6. **Did you feel that the teaching method employed in the lessons was appropriate?**
                        Yes   No

7. If no, what do you think could be improved/ altered?

      a. the music /excerpts used
      b. time management in the lessons
      c. fewer people in the classroom
      d. teacher's attitude
      e. Other (please specify)

      _____

      _____

      _____

      _____

8. How would you rate the level of difficulty of the lessons overall?

      a. Very difficult
      b.Difficult
      c.Neither difficult nor easy
      d.Easy
      e.Very easy

f. Did you feel that solfège skills have contributed in improving your level of dealing with music overall?

                    Yes   No

g. Do you have any further suggestions/comments regarding solfège lessons?

      _____

      _____

      _____

      _____

      _____

      _____

      _____

      _____

Thank you for your time

Appendix 2

Memorisation stimuli

Tonal stimuli – Pilot

T1 for piano:

T1 for violin:

T2 for clarinet:

T3 for flute:

Atonal stimuli – Pilot

A1 for piano:

A2 for soprano:

Tonal stimuli – Main study

T1 for euphonium:

T2 for violoncello:

T3 for piano:

Atonal stimuli – Main study

A1 for alto:

A2 for viola:

A3 for piano:

Appendix 3

Task questionnaires

Tonal pieces questionnaire –
Non-solfège participants

QUESTIONNAIRE

Please complete the following questionnaire by circling the appropriate answer or by writing it in the space provided. All responses will be treated anonymously and confidentially.

1. **Age:** _____

2. **Sex:**   Male   Female

3. **Degree Programme:** _____

4. **Year of study:** _____

5. **Instrument:** _____

6. **Instrument level:** Grade _____

7. **How many years of instrumental lessons have you had to date?** _____

8. **How many years in total have you been playing the instrument?** _____

9. **If you play another instrument/instruments in addition to the one you have been tested on, please fill the table below as appropriate:**

Other instrument(s)	Years of playing	Years of lessons

10. **How many years of music education have you had in total?** _____

11. **In which country did your main music education take place?** _____

12. **Do you have absolute pitch?**
    a. Yes
    b. No
    c. I don't know

13. How often do you memorise music for performance?
   a. Always/ almost always
   b. Often
   c. Occasionally
   d. Never/ almost never

14. How important do you consider memorisation skills to be for a musician?
   a. Extremely important
   b. Important
   c. Neither important nor unimportant
   d. Unimportant
   e. Extremely unimportant

15. How would you rate your memorisation skills:
   a. Extremely good
   b. Good
   c. Neither good nor poor
   d. Poor
   e. Extremely poor

16. Have you ever been taught how to memorise music?   Yes   No

17. If yes, who was it that taught you?
   a. School teacher
   b. Private teacher
   c. Conservatory teacher
   d. University teacher
   e. Friend/ colleague
   f. Other (please specify) _____

18. Do you employ any particular strategies for memorising music (e.g. repeating from beginning to end, dividing in sections etc)?
   a. Yes
   b. No
   c. I don't know

19. If yes, please select from the list below the strategy/strategies which you employ most often:
   a.   Repeating many times from the beginning to end
   b.   Dividing in sections (chunking) and learning them separately
   c.   Studying the music without the instrument, determining its structural elements
   d.   Studying the music without the instrument, going through the piece mentally
   e.   Other (please specify) _____
        _____

20. Please rate the difficulty of the items presented below by ticking the box which best reflects your opinion:

	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
Memorising music in general is					
The piece I have just been asked to memorise was, from a technical point of view					
The piece I have just been asked to memorise was, from a memorization point of view					

21. Did you feel you had enough time to memorise the piece?
   a. I had plenty of time
   b. I had just about enough time
   c. I needed a bit more time
   d. I needed a lot more time

22. Do you think you would memorise the piece in a different way if you weren't told that you will be asked to perform it again later?
                 Yes   No

23. If yes, what would you do differently?
   _____
   _____
   _____

24. Did the fact that you were being recorded affect your overall performance?
                 Yes   No

25. If yes, please circle the answer that best applies to you:
   a. It affected me very little
   b. It made me self-conscious but did not critically affect my performance
   c. It made me self-conscious and was a source of distraction from the task
   d. It impaired my performance severely
   e. Other (please specify) _____

   _____

Thank you very much

Tonal pieces questionnaire –
Solfège participants

QUESTIONNAIRE

Please complete the following questionnaire by circling the appropriate answer or by writing it in the space provided. All responses will be treated anonymously and confidentially.

1. **Age:** ＿＿＿

2. **Sex:**   Male   Female

3. **Degree Programme:** ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

4. **Year of study:** ＿＿＿＿＿＿＿

5. **Instrument:** ＿＿＿＿＿＿＿＿＿＿＿＿＿＿

6. **Instrument level:** Grade ＿＿＿＿＿＿＿

7. **How many years of instrumental lessons have you had to date?** ＿＿＿＿＿＿

8. **How many years in total have you been playing the instrument?** ＿＿＿＿＿＿

9. **If you play another instrument/instruments in addition to the one you have been tested on, please fill the table below as appropriate:**

Other instrument(s)	Years of playing	Years of lessons

10. **How many years of music education have you had in total?** ＿＿＿＿＿＿

11. **In which country did your main music education take place?** ＿＿＿＿＿＿＿

12. **Do you have absolute pitch?**
    a. Yes
    b. No
    c. I don't know

13. When were you taught solfège for the first time? Age _____

14. Had you had any music lessons before you started using solfège? Yes   No

15. If yes, for how many years? _____

16. Do you automatically think of note names (e.g. do, re, mi, fa etc) when reading music?
   a. Yes
   b. No
   c. Sometimes

17. Do you automatically think of note names (e.g. do, re, mi, fa etc) when playing/singing music?
   a. Yes
   b. No
   c. Sometimes

18. Do you normally use solfège in your everyday music practice?
   a. Always/ almost always
   b. Often
   c. Occasionally
   d. Never/ almost never

19. Do you normally use solfège when memorising music?
   a. Always/ almost always
   b. Often
   c. Occasionally
   d. Never/ almost never

20. If your answer to the previous question was b, c, or d, which are the circumstances that make you use solfège when memorising music? (circle as many as apply to you)
   f.   I use solfège when I find music more difficult than average
   g.   I use solfège when I find music easier than average
   h.   I use solfège when there is time pressure
   i.   I use solfège when there is no time pressure
   j.   Other (please specify)_____

   _____

21. How often do you memorise music for performance?
   a. Always/ almost always
   b. Often
   c. Occasionally
   d. Never/ almost never

22. How important do you consider memorisation skills to be for a musician?
    a. Extremely important
    b. Important
    c. Neither important nor unimportant
    d. Unimportant
    e. Extremely unimportant

23. How would you rate your memorisation skills:
    k. Extremely good
    l. Good
    m. Neither good nor poor
    n. Poor
    o. Extremely poor

24. Have you ever been taught how to memorise music?   Yes   No

25. If yes, who was it that taught you?
    a. School teacher
    b. Private teacher
    c. Conservatory teacher
    d. University teacher
    e. Friend/ colleague
    f. Other (please specify)_____

26. Do you employ any particular strategies for memorising music (e.g. repeating from beginning to end, dividing in sections etc)?
    a. Yes
    b. No
    c. I don't know

27. If yes, please select from the list below the strategy/strategies which you employ most often:
    p. Repeating many times from the beginning to end
    q. Dividing in sections (chunking) and learning them separately
    r. Studying the music without the instrument, determining its structural elements
    s. Studying the music without the instrument, going through the piece mentally
    t. Other (please specify)   _____

_____

_____

28. Please rate the difficulty of the items presented below by ticking the box which best reflects your opinion:

	Very difficult	Difficult	Neither difficult nor easy	Easy	Very easy
Memorising music in general is					
The piece I have just been asked to memorise was, from a technical point of view					
The piece I have just been asked to memorise was, from a memorisation point of view					

29. How did the fact that you were specifically asked to use solfège affect your memorisation?
    u. It had no effect on my memorization
    v. It had a negative effect on my memorisation
    w. It had a positive effect on my memorisation
    x. Other (please specify) _____

30. If your answer to the previous question was b) or c), please explain your response further: _____

_____

31. Did you feel you had enough time to memorise the piece?
    a. I had plenty of time
    b. I had just about enough time
    c. I needed a bit more time
    d. I needed a lot more time

32. Do you think you would memorise the piece in a different way if you weren't told that you will be asked to perform it again later?   Yes   No

33. If yes, what would you do differently?
_____

_____

34. Did the fact that you were being recorded affect your overall performance?
                     Yes   No

35. If yes, please circle the answer that best applies to you:
    y. It affected me very little
    z. It made me self-conscious but did not critically affect my performance
    aa. It made me self-conscious and was a source of distraction from the task
    bb. It impaired my performance severely
    cc. Other (please specify) _____

Thank you very much

Additional questions for the atonal pieces – All participants

1. **How familiar are you with atonal music?**
    a. Very familiar
    b. Familiar
    c. Neither familiar nor unfamiliar
    d. Unfamiliar
    e. Extremely unfamiliar

2. **How often do you practise/perform atonal music?**
    a. Extremely often/ often
    b. Sometimes
    c. Occasionally
    d. Never/ almost never

3. **Have you had any systematic experience of practising/performing/learning atonal music?**
    Yes   No

4. **If yes, what was the nature of this experience?**
    a. Seminar(s)
    b. Masterclass(es)
    c. Studying an atonal piece with an instrumental/vocal tutor
    d. Playing an atonal piece in an orchestra/ensemble
    e. Other (please specify)_____

5. **When you are asked to memorise an atonal piece of music, do you do it in the same way as you memorise tonal music?**
    Yes   No

6. **If no, what do you do differently?**
    _____
    _____

Experiment 3 questionnaire –
All participants

QUESTIONNAIRE

Please complete the following questionnaire by circling the appropriate answer or by writing it in the space provided. All responses will be treated anonymously and confidentially.

1. **Participant number:** _____

2. **Age:** _____

3. **Sex:**   Male   Female

4. **Dominant hand:**   Right   Left   Ambidextrous   I don't know

5. **Do you have absolute pitch?**   Yes   No   I don't know

6. **Degree Programme & Year / Position:** _____

7. **First study instrument:** _____

8. **How many years of instrumental lessons have you had to date?** _____

9. **How many years in total have you been playing the instrument?** _____

10. **If you play another instrument/instruments in addition to your first study instrument, please fill the table below as appropriate:**

Other instrument(s)	Years of playing	Years of lessons

Thank you for your time

Appendix 4

Transcriptions and scoring

Sample practice transcriptions

Participant no. 31 (soprano, non-solfège), Tonal piece:

PV 1:30. B.1 (piano), plays 2t, sings (aaa). Same in b.2, then together. Then b.3 vo + pn, only vo and corrects with pn, b.4 vo + pn, b.5 pn, then vo, 6-8 pn. Then vo only (corrects with pn) 1-4, concentrates on passages, repeats, corrects with pn, for 2'. Then the same system of studying for b.5-8. (4:30). Plays once all on pno. Sings start-end from score, then by heart (perhaps periodically looks), concentrates on links (e.g. b.4-5). Same system until the end, sings top-end starting with chord in the beginning.

Participant no. 91 (violin, solfège), Atonal piece:

Once rhythmic solfège top-end, then PV quite good 40''. Replays carefully top-end – first tries to play singing solfège but abandons after first few notes), plays 1-8, rhythm solfège 1-4 several times and then plays 1-4 several times (probably starts memorising – can be heard walking away while playing), tries to sing simultaneously. Then plays b.5 carefully several times, 5-6 several t., 7-8, 1-8 several times carefully, (5'), sometimes starting from different bars, possibly by heart and sometimes tries to start with solfège.

Participant no. 44 (flute, solfège), Atonal piece:

PV 40'' – restart at b.5. Plays 1-2, tries by heart, cannot, does solfège. 1-4 several times. Then does solfège 5-6, very well (checks pitch with flute), plays 5-6, then 1-6 + 7, link 6-8, starts from 5, gets confused and remembers it with solfège (cannot tell whether solfège is from memory though, it helps her in some way). Plays again 5-8. (5'). Plays top-end, begins solfège in b.1, plays again top-end and again solfège b.1. Does solfège top-end.

Transcription key

PV	Prima vista (sight-reading)
1'	1 minute
1''	1 second
b.1	bar 1
b.1-2	bars 1-2
1	bar 1
1-2	bars 1-2
t.	times
vo	voice
pn	piano
.(5').	The fifth minute of the practice session.
, (5'),	meanwhile, the 5th minute of the practice session passes

Scoring method

Pitch accuracy (per notes played)

Type of error		Penalty
Wrong note		- 1/12 for each semitone difference
Octave difference (single notes)		- 2
Octave difference (2 or more notes)		see Transposition
Transposition (motifs/phrases)	First note	- 1/12 for each semitone difference
	Remaining notes	- P/2
Repeated erroneous motif		- 1/12 for each semitone difference for the 1st time only
Notes of a motif in wrong order	Tonal piece	- P/2
	Atonal piece	- 1/12 for each semitone difference
Inverted repeated motif		- 1/12 for each semitone difference for the first time in each bar only
Doubled wrong note		- 1/12 for each semitone difference only for the 1st note
Missing doubled notes		Impose penalty only in Fluency/Structural accuracy, but not in Pitch accuracy
Extra notes (within metre)		- 1/12 for each semitone difference from the original note
Playing phrase from another part of the piece in order to fill a gap in rhythm		Count as missing notes only, but not as extra
Pianists: Switching melody from left to right hand or the opposite		Ignore octave difference, impose Transposition penalty

Correct note = + 1 pt

Rhythm accuracy (per notes played)

Type of error	Penalty
Wrong value in 4/4 metre	- 0.25 for each semiquaver difference
Wrong value in 6/8 metre	- 1/6 for each semiquaver difference up to a crotchet, -1 for each dotted crotchet difference.
Correct values within motif – wrong positions	- P/2
Playing phrase from another part of the piece in order to fill a gap in rhythm	Count as missing notes only, but not as extra
Extra notes	Impose penalty only in Fluency/ Structural accuracy, but not in Rhythm accuracy

Correct note = + 1 pt

Structural accuracy (per total bars)

Type of error		Penalty
Correct bar(s), wrong position		- 1 for the first bar in the sequence
'Filling' metre with wrong notes	Tonal piece	- 0.75
	Atonal piece	- 1
Less beats in a bar in 4/4 metre		- 0.25 for each semiquaver difference
Less beats in a bar in 4/4 metre		- 0.25 for semiquaver differences - 0.5 for quaver differences - 0.75 for crotchet differences - 1 for dotted crotchet differences
Wrong harmony – Tonal piece		- 0.1
Extra bars		- 0.25 each
Extra trial beats		- P/2

Correct bar, correct position = + 2 pts

Fluency (per beats played)

Type of error	Penalty
Appoggiatura	- 0.25
Out of tune, less than 5 notes total	- 0.25 each
Out of tune, 5 or more notes in total	- 1 per phrase
Instant correction of one wrong note	- 0.25
Correction of one note with pause	- 0.5
Correction 2 notes	- 0.5
Correction 3 notes or more	- 0.75
Delay less than 0.5 beat	- 0.25
Delay less than 1 beat	- 0. 5
Delay between 1–2.5 beats	-1
Delay between or equal with 2.5–5 beats	-2
Delay more than 5 beats	- $(N_{beats} - 1)/2$
Ritenuto / Acccelerando	- $0.25 \times N_{beats}$
Extra notes	- 0.75
Repeat note / motif	- 0.75 per note/motif
Restart	- 0.75
Tempo change	- 2 per change
General looseness of rhythm	- 1
Talk / Laugh	- 2
Pianists: use of pedal	- 0.5 for the whole piece

Correct beat = + 1 pt

Appendix 5

Δ scores regression results

Regression analysis results for Δ note span scores

Δ_{AC1} (N = 19)		B	SE B[a]	β	$R^2 = .00$ ns
	Constant	- .33	.43		
	Solfège knowledge	.13	.75	.04*	

Δ_{AC2} (N = 22)		B	SE B[a]	β	$R^2 = .13$ ns
	Constant	- .60	.68		
	Solfège knowledge	1.43	.87	.35*	

Δ_{AC3} (N = 22)		B	SE B[a]	β	$R^2 = .00$ ns
	Constant	- .11	.89		
	Solfège knowledge	- .12	.93	- .03*	

Δ_{C31} (N = 16)		B	SE B[a]	β	$R^2 = .10$ ns
	Constant	75.44	7.06		
	Solfège knowledge	10.89	7.18	.32*	

Δ_{C32} (N = 19)		B	SE B[a]	β	$R^2 = .16$ ns
	Constant	80.67	2.48		
	Solfège knowledge	5.78	2.82	.41*	

Δ_{C12} (N = 14)		B	SE B[a]	β	$R^2 = .12$ ns
	Constant	.60	.38		
	Solfège knowledge	- 1.16	.72	- .35*	

Regression analysis results for Δ distraction task scores

ΔD_{AB1} (N = 25)		B	SE B[a]	β	$R^2 = .01$ ns
	Constant	1.94	6.17		
	Solfège knowledge	- 3.67	7.20	- .11*	

ΔD_{AB2} (N = 25)		B	SE B[a]	β	$R^2 = .03$ ns
	Constant	- 6.18	3.67		
	Solfège knowledge	5.51	6.32	.16*	

ΔD_{AB3} (N = 25)		B	SE B[a]	β	$R^2 = .10$ ns
	Constant	- 1.57	3.32		
	Solfège knowledge	8.51	5.52	.31*	

ΔD_{B31} (N = 17)		B	SE B[a]	β	$R^2 = .00$ ns
	Constant	.41	4.72		
	Solfège knowledge	.05	6.63	- .00*	

ΔD_{B32} (N = 18)		B	SE B[a]	β	$R^2 = .28$ p < .05
	Constant	- 3.97	3.49		
	Solfège knowledge	10.46	- 0.11	- .53**	

ΔD_{B12} (N = 18)		B	SE B[a]	β	$R^2 = .00$ ns
	Constant	.56	5.36		
	Solfège knowledge	2.16	6.82	.07	

Regression analysis results for Δ verbal span scores

Δ_{B31} (N = 17)		B	SE B[a]	β	$R^2 = .00$ ns
	Constant	- 1.43	7.83		
	Solfège knowledge	- 4.64	13.61	- .09*	

Δ_{B32} (N = 18)		B	SE B[a]	β	$R^2 = .20$ ns
	Constant	- 8.63	3.86		
	Solfège knowledge	20.44	10.55	.45*	

Δ_{B12} (N = 18)		B	SE B[a]	β	$R^2 = .07$ ns
	Constant	9.86	11.57		
	Solfège knowledge	14.64	13.30	.27*	