Preview, perception and motor skill in piano sight-reading

Submitted for the degree of PhD

by

David Evans

Department of Music

University of Sheffield

March 2008

Abstract

Ten skilled and eleven less skilled sight-readers, advanced adult pianists, undertook two sets of studies. First, controlled preview experiments measured dependence of maximised sight-reading tempo on preview size with monophonic, two-part and four-part notation. Secondly, monophonic experiments measured isolated, sight-reading related perceptual and motor sub-skills: a transcription based error detection task and a test of visually unmonitored, unrehearsed output. All experiments employed tonally coherent and incoherent materials, with the aim of testing the theory that the ability of skilled readers lies in their use of larger preview than less skilled readers, courtesy of their greater sensitivity to musical structure.

The skilled group sight-read consistently faster than the less skilled group, achieving larger effective preview with monophonic and four-part, but crucially not with two-part materials. Extra preview use with the former materials was found to be a source of only small gains, the evidence overall indicating skilled readers' faster performance to have been primarily dependent on a more efficient processing of smaller preview amounts than less skilled readers. Both skill groups demonstrated similar, limited tempo responses to the structural distinction in experimental materials, with no structural effect on preview for skilled readers. These results suggest, therefore, that the skilled group's superior performance was primarily due to perceptuo-motor factors.

This finding is confirmed by the skilled group's faster performance on the two sub-skill studies. On the perceptual study, both groups display similar patterns of response and sensitivity to structure. In terms of motor skill, compared to less skilled readers, skilled readers are either better at unrehearsed output, non-visually monitored performance, or both. Finally, individual participant data suggest sight-reading to be a complex combination of skills: many participants show significant variation in performance across the studies, and there is evidence for a number of different factors limiting skill development amongst less skilled readers.

Acknowledgements

In relation to the research itself, I would first of all like to thank Prof. Eric Clarke for being the very best of supervisors. Also, I am profoundly grateful to the experimental participants, who were so generous with their time, and who carried out the work with such commitment and enthusiasm. Finally, thanks are due to Dr. Colin Peters for help with the experimental materials, and Prof. Phil Prescott for advice concerning the statistical analysis.

Particular thanks are also due to my father, my wife and children for all their loving support, and for giving me the time and space to carry out this work. The same goes for my mother, who sadly died in the middle of the research process. This thesis is dedicated to her.

Table of contents

List of tables

List of figures

Chapter 1

Introduction

1 Introduction

Within a musical context, the term 'sight-reading' can be interpreted in various ways. Perhaps in its purest incarnation it can be defined as the performance in real-time of previously unseen musical notation, and this is the sense in which the term is to be understood in relation to this thesis. As such, it is a transcription task belonging to the same genre of skills as touch-typing and the reading aloud of text, where abstract external representations (music notation/language text) are translated into their corresponding motor counterparts to produce tones from musical instruments, depress typewriter keys or to manipulate the vocal apparatus. But unlike these other transcriptional activities the term also encompasses the idea of a certain level of pre-reading prior to the performance of the entire task. For example, music exam candidates typically have a limited period within which to prepare their sight-reading performances. What sight-reading represents in this case is the idea that the material has not been subjected to any rigorous and sustained rehearsal. Furthermore, an orchestral conductor mentally reading through a musical score without producing any motor output at all could also be described as engaging in sight-reading. Additional variation in meaning is found when the nature of the performance output is considered. The task may entail an attempt at absolute accuracy of transcription, or merely an apt approximation, for example, the improvised 'bluffing' of professional accompanists. All of these manifestations of sight-reading are valuable focuses of research, but the semantic ambiguity must be carefully borne in mind.

An obvious, necessary requirement for music sight-reading is a certain expertise in the performance of the particular musical instrument itself. Unlike skilled touch-typing, which exists essentially as a transcription task, music performance and the vocalising of text are not necessarily defined by a real-time visuo-motor transcriptional element. Actors and solo musicians typically learn their material by rote, enabling performance to be carried out with recourse to only memorised representations. Expertise at rehearsed performance, though, does not necessarily correlate with superior sight-reading ability. There does appear to be a large variability in sight-reading skill amongst pianists at all levels (Lehmann & Ericsson, 1993). This seems to be less so for single melody line instruments, for example, the clarinet and trumpet (McPherson, 1995). Making anecdotal reference on this point, the pianist and teacher Kendall

Taylor writes that 'players of orchestral instruments, who read one stave at time, are usually fluent readers' (Taylor, 1981, p. 151). This difference is no doubt due, in part, to the smaller quantity of transcriptional throughput that is required with such instruments, with less demand being placed upon limited processing resources. However, much less is documented about sight-reading skill on instruments other than the piano. It has been the piano that has been the main focus of research interest, partly, it would seem, because of interest in the extra complexity involved in piano-based reading, but also because of the ease with which accurate data of piano performance can be obtained using MIDI technology.

The demands of successful music sight-reading are considerably greater than for a transcriptional task like touch-typing, involving not only the production of correct pitches on the instrument, but also their performance to a defined schedule of timing, together with varied articulation in a stylistic interpretation requiring both intellectual grasping of musical structure and artistic/emotional sensibility. For the piano, the particular focus of this thesis, choices also need to be made concerning the fingering that will best perform the notation, and the complex and subtle adjustments of finger, hand and arm to provide attack to the keys giving the desired expressive effect. Also with the piano, the sight-reader is typically required to read two staves of music notation (usually one for each hand) and to perform up to five notes (or possibly more) simultaneously per hand.

For the psychologist, the intricacy of the task together with the apparently wide skill differences prevalent amongst musicians of equivalent performing standards makes the study of piano sight-reading a potentially rich resource of insight into complex perceptual, cognitive and motor processes, together with the development of expertise. However, such insights have a potential practical relevance to musicians and educators too, in terms of how they might inform skill training. Although the ability to read well is not strictly necessary for the solo-performing musician, it is of immense benefit, enabling new repertoire to be explored with ease. For session musicians and busy accompanists who have little time to prepare performances, it is essential. Above all, printed notation is the principal means of disseminating music, and to struggle with such notation is indeed to be cut off from the *lingua franca*. In view of this, it is common to find researchers and authors discussing the implications of empirical findings for pedagogical practice (Sloboda, 1978b; Lehmann & McArthur, 2002; Lee, 2004; Thompson & Lehmann, 2004; Lehmann, Sloboda & Woody, 2007).

The usefulness of such research-based advice to skill development is clearly dependent upon the validity of the underlying research findings and their interpretation. The principal researchers within the domain appear to be quite upbeat about the current state of knowledge. For example, in a recent review, Lehmann and McArthur write that 'the various research strands enable us to develop a cognitive model of sight-reading that is both consistent with research findings and of practical interest to music practitioners' (Lehmann & McArthur, 2002, p. 144). This model dominates current thinking within the domain, and provides a particular account of sight-reading skill difference amongst proficient musicians. The account will be described in detail in Chapter 2, but essentially it holds that through their greater experience at the task, skilled readers become considerably more sensitive to musical structure within the score than less skilled readers, something that provides them with highly effective task-specific memory mechanisms. Grasping musical meaning enables skilled readers to store notes more efficiently in memory and to process them at a faster rate. Less skilled readers, on the other hand, have to rely more on processing notes as individual, unrelated entities. The account's proponents present a range of research findings in support of these ideas. For example, it is claimed that skilled readers scan notation more efficiently than less skilled readers, perceive more notes within individual eye-fixations, use larger perceptual groupings, and gather information from further ahead in the score with which to plan their motor responses. The account would not deny that skilled readers might also develop more effective basic perceptual and motor processes in relation to the task, and become more proficient at other cognitive strategies unrelated to specifically musical content, for example, using intervallic-based information. However, such mechanisms are considered thoroughly incapable of providing the quantity and speed of note processing necessary for skilled levels of reading.

There would seem on the surface to be much to commend this account of sight-reading ability, and the ideas upon which it is based have certainly become widely accepted. However, a detailed examination of the associated research literature (fully documented in Chapter 2) indicates that there may be rather less empirical support for it at present than its proponents have considered. For example, whilst some research has provided evidence of skilled readers processing material further ahead in the score than less skilled readers, other work has found them to differ little in this regard. Also, there is no clear evidence that skilled readers are more sensitive to structure in notation than less skilled readers, meaning that where skilled readers are found to process notes in larger units, this might be due to other factors. Furthermore, no research has been carried out that quantifies the extent of the sight-reading performance gains that accrue from structural perception. It could therefore be that the structure-based account has been unduly pessimistic about the potential of more basic information processing mechanisms to support a skilled level of performance.

Clearly, the lack of empirical support does not necessarily indicate that a primarily structure-based account of reading skill is inappropriate. It does, however, mean that the role of non-structural factors, for example, elemental perceptual and motor skill cannot, at least at present, be dismissed so readily in terms of their potential as significant sources of skill variation. It would seem, therefore, that the proponents of the structure-based account might have been a little presumptuous in their theorising. Perhaps this would be of only limited concern if it were confined to merely academic circles. However, the account has been presented recently in a number of publications aimed at music professionals and students, with virtually no attention paid to the possibility of alternative interpretations of the available evidence (Lehmann & McArthur, 2002; Thompson & Lehmann, 2004; Lehmann *et al.*, 2007). In view of the limited research evidence available, the wisdom of this would seem to be questionable.

The review of the literature undertaken in Chapter 2 indicates that the available research data on sight-reading is inadequate to support any confident theorising at present. The approach to research has been too piecemeal, meaning that there are simply too many gaps in our knowledge, even about foundational issues, that need to be filled. It would seem to be a priority, therefore, for research to begin working towards developing a more complete understanding of the most basic elements of skill at the task, to provide a securer foundation both for theory development and for future research to build upon. This is the principal driving force behind the research direction of this thesis, which has focused upon gaining a clearer understanding of three fundamental areas of sight-reading skill.

Firstly, work has been carried out to try and gain more detailed knowledge about the perceptual requirements of musicians with regard to notation during sight-reading i.e. how far skilled and less skilled sight-readers need to look ahead in the music to support their performances (Chapters 4, 5, 6, 7, 8 and 9). It was mentioned earlier that research findings are not as clear-cut about skilled readers use of larger 'preview' as supporters of the structure-based account seem to have considered, and so it was important for work to be undertaken to clarify this matter. Previous preview-related studies have tended to focus upon a single notational complexity at a single tempo, providing what may only be a very partial insight into skill. To try and gain a broader understanding, the research in this thesis has explored both simple and more complex musical materials at different performance speeds. Secondly, an investigation has been undertaken into the speed and pattern of perceptual activity isolated from motor performance (Chapters 10 and 12). Finally, a study has explored sight-reading related motor ability, isolated from normal notational input (Chapters 11 and 12). As well as providing some foundational understanding of the specific sub-skills themselves, the findings of these experiments have the potential to inform the results of the complete task undertaken in the preview studies offering insight, for example, into the relative influence of perceptual and motor skill in explaining sight-reading ability. More detailed reasons for the particular choice of research areas is provided in Chapter 2 (the literature review) and an introduction to, and rationale for, the specific methodologies and technologies employed is found in Chapter 3.

Although the primary reason for the research directions chosen was to gain more detailed knowledge about foundational aspects of sight-reading skill, an advantage of the particular studies undertaken is that they also offer the opportunity of a clearer understanding of the relative importance of sensitivity to musical structure and basic perceptual and motor factors in determining the abilities of skilled and less skilled readers. There is a strong emphasis within the thesis on analysing the research findings in relation to these issues.

Chapter 2

Research into sight-reading: a critical view of the literature

2 Research into sight-reading: a critical review of the literature

2.1 Introduction

The study of music sight-reading has been carried out principally within the fields of psychology and music education. Within psychology, research has been mainly concerned with exploring the perceptual and cognitive processes involved in task performance, particularly in the context of seeking to explain variation in ability. Within educational research, study has typically been carried out into the effects of specific pedagogical regimens (e.g. Streckfuss, 1984; Kostka, 2000), with little attempt to isolate the specific roles of perceptual and cognitive factors. Therefore, whilst being able to offer potentially valuable skill-related insights, such research has been limited in its ability to provide any detailed dissection and explanation of the variation in sight-reading skill examined. Because of this, the review will focus principally upon the core psychological literature, considering educational research only where findings are felt to be of relevance to the particular issue being considered. This review of the literature is organised under the following headings, which encompass the range of potentially causative factors upon skill that have been the focus of research attention.

- Input and memory processing mechanisms:
    1. Visual perception and processing: eye-movement and perceptual research
    2. Musical structure and sight-reading skill
    3. Auditory representations

- Output mechanisms: psychomotor skill

- Research within the expertise paradigm: the role of practice in skill determination

- The measurement of sight-reading ability: a discussion of concepts and methods

- Conclusion and areas needing research

Before turning to a detailed discussion of the research literature, an important methodological issue needs to be briefly addressed. A problem with a number of pieces of research that will be considered in this review is the lack of control over the technical instrumental skill of participants. Sight-reading ability clearly has dependencies upon rehearsed performance ability - skilled, experienced instrumentalists who are poor sight-readers compared with their peers will nearly always be better at reading than novice instrumentalists who read well for their limited level of technical skill. Therefore, technical skill needs to be sufficiently controlled so that factors pertaining specifically to sight-reading skill can be isolated. Much of the early research into sight-reading carried out in the 1930s and 1940s, for example Weaver (1943) and Bean (1938), did not include such experimental controls, focusing instead simply upon experienced and less experienced musicians. Because of this, whilst there is much of interest and value in these investigations, they are of limited use in achieving a detailed understanding of the factors associated with sight-reading ability itself.

The issue began to be addressed with the work of Sloboda (1974), and since then research has generally concentrated its efforts upon investigating the sight-reading ability of skilled instrumentalists. However, a very few studies, for example Truitt, Clifton, Pollatsek and Rayner (1997), have continued without controls upon technical skill, and there have also been various pieces of research undertaken into the perceptual skills of musicians and non-musicians (Sloboda, 1976a, 1978a; Halpern & Bower, 1982). Reviewers of the literature have not always taken this range of participant make-ups and controls sufficiently into account when drawing conclusions about sight-reading ability, and it is not uncommon to see an inappropriate mixing and matching of empirical findings from studies involving non-equivalent groups of participants. To take one example amongst many, research data obtained from 'skilled' and 'novice' instrumentalists by Furneaux and Land (1999) are cited by Lehmann and McArthur (2002) as relating to 'good' and 'less skilled' readers. Such conceptual imprecision is clearly unsatisfactory for scientific discourse, and the achieving of an agreed and detailed terminology to describe the different types of participants that studies have employed would seem an urgent priority for the domain. This thesis will attempt to identify unambiguously the types of participants used within individual pieces of research to hopefully avoid any misrepresentation of study findings. Particularly, the terms 'skilled reader' and 'less skilled reader' will be used to refer to skilled musicians only.

Although recent research has controlled more for technical skill, it must be borne in mind that any insights obtained into sight-reading ability are only as dependable as the quality of the controls employed. A case in point is the work of Waters, Townsend and Underwood (1998a). These researchers chose participants aged 18-25 years who were required to have had piano training for at least 5 years, something that would seem to allow for some considerable variation in instrumental performance skill level that might result in a clouding of specific sight-reading related effects.

2.2 Input and memory processing mechanisms

2.2.1 Visual perception and processing

2.2.1.1 Introduction

The majority of studies in music sight-reading have researched input related factors, either in the context of complete sight-reading performance or else isolated from normal motor output. Despite the fact that the former type of research involves a significant motor component, researchers have typically attributed input or memory processing related explanations to variation in the measures obtained for skilled and less skilled readers, failing to give sufficient consideration to the possibility of motor influence. The tendency has been to view skill at rehearsed performance as somehow sufficient evidence of the possession of motor skills appropriate to skilled reading. Only recently, it would seem, have music psychology researchers begun to consider that rehearsed movements and the short-notice, 'online' movements involved in tasks like sight-reading and improvisation (Thompson & Lehmann, 2004) might have different underlying skill requirements. This issue will be examined in more detail, later, in the section on output mechanisms.

As discussed briefly in Chapter 1, psychological study into the origins of sight-reading ability is dominated by a single account, which holds that skilled readers owe their ability to being more sensitive to meaningful musical structure within the score than less skilled readers, an explanation that I will from now on refer to as 'the patterning account'. Since its formulation by Sloboda (1974) in the 1970s, this account has to my knowledge never been subjected to any substantial published critique, and no other theory has been seriously proposed as an alternative. Indeed, within recent reviews of the literature like Lehmann and McArthur (2002) and Lehmann *et al.* (2007), the authority with which it is expounded provides clear evidence of its now widespread acceptance. According to these authors, the patterning account is not only supported by dedicated research into the role of musical structure within sight-reading, it is also consistent with the findings of more general work into visual input processes - eye-movement and perceptual studies. In view of this, the account would appear to provide a useful framework within which to organise and coordinate an exploration of the literature relating to these three different aspects of skill.

I will begin this exploration of visual input and memory processes with an extended summary of the patterning account itself, to provide a context for subsequent discussion. In the light of the summary, I will then examine, first of all, eye-movement research, secondly visual perception research, and finally work into musical patterning itself, and from this attempt to gauge the extent of support currently available for the patterning account, as well as determining whether any alternative interpretations of the evidence might be indicated.

2.2.1.2 Summary of the patterning account

According to the proponents of the patterning account, skilled music sight-reading demands large preview (the extent of perception beyond the note currently being performed) both to facilitate a fast encoding of notation (Lehmann *et al.*, 2007) and fluent motor planning (Thompson & Lehmann, 2004). Such preview is considered beyond the capacity of normal information processing mechanisms to achieve, in which notes are perceived as unrelated units and only processed individually, or in small groups at a time, from the score. Such processing is viewed as memory inefficient both in terms of speed and size of short-term storage, resulting in slower and disjointed motor output. To attain the quantity and speed of throughput for skilled reading, experts have to 'circumvent limitations of the human information-processing system' (Lehmann *et al.*, 2007, p. 114). It is proposed that they do this by perceiving and processing larger quantities of notes in meaningful, more memory efficient groups or chunks, a consequence of having prior knowledge and sensitivity to the musical grammar and style of the material being performed. 'The available evidence suggests that readers require preview of structural units within a text if they are to organize fluent and rapid performance' (Sloboda, 1985, p. 71).

There are three ways in which it is considered that these larger chunks are attained: "by intelligent anticipation…and problem solving or by creating long-term working memory structures" (Lehmann *et al.*, 2007, p. 114). Long-term working memory structures are regularly met patterns of notation within a particular grammar or style, for example scales and broken chords, which are stored in long-term memory. According to the long-term working-memory theory of Ericsson and Kintsch (1995), skilled readers are considered to have a high-speed connection to such patterns, something that considerably lightens the load upon short-term and working storage, and leads to highly automated motor output. Lehmann *et al.* (2007) demonstrate this effectively by presenting the reader with two strings of text, one a randomised version of the other: "g-s-n-i-i-g-d-h-a-t-e-r" and "s-i-g-h-t-r-e-a-d-i-n-g". For confident readers of the English language, the latter can be stored quickly and efficiently in memory as a single, meaningful chunk. The former, however, requires storage to be broken down into smaller units, making the whole string take longer to process, and be less efficiently and securely stored in short-term memory. Sloboda (1985) provides another example that graphically illustrates how chunking in this manner influences transcriptional output in the related task of touch-typing. He quotes from work by Shaffer (1976) showing that when skilled typists were required to type text made up of words with randomised letter order, their speed of output was reduced from normal levels (8 characters per second) down nearly to their basic reaction level response with individually presented single letters (2 characters per second).

The other two methods proposed for increasing the throughput of notation are closely related to each other. Because of its fundamentally patterned nature, music necessarily has predictable elements, meaning that notes do not always require detailed individual processing, but can be inferred, not necessarily consciously, from their context. 'Problem solving' refers to the production of structurally inferred output during the performance of an incompletely processed group of notes (Lehmann & McArthur, 2002), whereas 'intelligent anticipation' involves the making of hypotheses about an oncoming section of notation, for example, in relation to the continuation of melodic sequences. Clearly, the larger the chunks that a sight-reader is using, the more musical context there is available, increasing the likelihood that anticipation and inference will be successful. Concerning monophonic reading, Sloboda writes that "the greater capacity of this store *(i.e. short-term memory)*, the greater is the opportunity for preview, and the greater is the opportunity for making reasonable predictions about subsequent notes. Given six notes to go on, success in predicting the next note is more likely than if there are only two notes" (Sloboda, 1978b, p. 12).

The perception of musical structure mediated in these three ways is viewed as essential to the explaining of skilled reading. Skilled readers do not read note-by-note but depend on seeing general patterns and reconstructing the details of the music using their prior knowledge. Less skilled readers' weaker performance, on the other hand, is primarily put down to their inability to implement these mechanisms, a consequence of their being less sensitive to musical structure within the score. As a result, they cannot bypass information processing limitations, as skilled readers do, and so are confined to smaller chunk sizes, and consequent slower perception of preview together with disjointed output. The patterning account has not traditionally given any credence to the idea that less skilled readers' lack of ability may be primarily the result of inferior perceptuo-motor functioning. As mentioned in the Chapter 1, its proponents do not consider that skilled and less-skilled readers will show no variation in such functioning, but the general view appears to have been that this is not a limiting factor upon skill; to achieve expertise, basic information-processing methods need be transcended, not improved. This is clearly seen in Sloboda's advice to less skilled sight-readers seeking to develop their abilities. 'Sight-reading performed note-by-note is unlikely to improve however frequently practiced. The music must be understood before it is played' (Sloboda, 1978b, p. 15). Recently, there has been a greater openness amongst supporters of the patterning account to the idea of motor ability being an important factor in explaining sight-reading skill difference (Thompson & Lehmann, 2004). However, such thinking is potentially undermining of their central thesis. If motor factors are seen as important in explaining sight-reading skill variation, it is inconsistent to continue to assume that the variation in performance measures that underlie the patterning account (for example the amount of preview employed by musicians) is necessarily driven by sensitivity to musical structure.

2.2.1.3 Eye-movement research

Proponents of the patterning account regard the findings of eye-movement research to be consistent with and reflective of their understanding of the perceptual roots of sight-reading skill difference. 'It has become clearly established that unskilled readers differ markedly from experienced readers with regard to their looking behaviour' (Thompson & Lehmann, 2004, p. 146). More specifically, 'better readers will scan (or parse) the page more efficiently than less-skilled readers and they require, 'shorter and fewer fixations to compare or encode material for execution because they are able to grasp more information in one fixation (Lehmann and McArthur, 2002, p. 138). Better readers are also considered to have larger eye-hand spans; that is, their eyes 'look further ahead of the point where they are currently playing' (Thompson & Lehmann, 2004, p. 146). Skilled readers 'do not fixate on all notes', whereas 'less proficient readers tend to focus on individual notes' (Lehmann & McArthur, 2002, p. 138). 'Less proficient readers search around for information and try to make sense of what they see, while efficient readers seem to know what to look for' (Lehmann & MacArthur, 2002, p. 138). The studies cited in relation to these quotes, which will be explored shortly, involved skilled and less-skilled sight-readers who were experienced musicians, and so were explicitly concerned with isolating factors specifically relating to sight-reading ability. Lehmann and MacArthur also state unequivocally that their references represent a summary of research findings. A detailed examination of the eye-movement literature, though, suggests a somewhat more varied picture to the one that has been painted.

There are principally three issues to be considered:

- Do skilled readers use shorter and fewer fixations than less skilled readers?
- Do skilled readers take in larger groups of notes in their fixations and have larger eye-hand spans? Do less skilled readers typically fixate on notes individually?
- Does the eye-movement behaviour of skilled readers show greater evidence of a planned strategy?

Before discussing these issues it is necessary to examine briefly the history of eye-movement research. Study in this area began with the work of Jacobsen (1941) and Weaver (1943) who employed rudimentary devices that could photograph the position and timing of eye fixations in relation to sight-read notation. There are some significant problems with this early work, though. To begin with, the technology was limited in terms of the accuracy with which the position of eye-fixations upon the score could be determined, and the rate at which sampling of movements could be achieved. Also, normal performance-related bodily activity, including the tendency of participants to regularly glance down at their hands during the reading task, inevitably led to contamination of the data. To attempt to minimise this effect, participants had to be trained to limit their musculoskeletal movement, and various physical restraints like headrests and bite-plates were employed to this end. This restraint upon movement, particularly the prevention of looking down at the hands (an activity natural to even skilled sight-readers (Banton, 1995)) raises significant doubts about the ecological validity of the research. A further point that needs to be mentioned is that because of the typical lack of controls made upon technical skill, little insight can be gained into how recorded patterns of eye-movement reflect factors specific to sight-reading ability itself. It was not until the work of Goolsby (1994a) that equipment was available sensitive enough to define eye position to within the size of an individual note, in the context of adequate sampling rates. However, even with this more recent technology, some concerns remain over the ecological validity of findings because of the issue of constraints upon bodily movement.

Considering research findings relating to eye fixations to begin with, the results of early research, for example Jacobsen (1941), are in agreement with Lehmann and MacArthur with regard to better readers using a shorter fixation length, but typically show them using more frequent, not fewer, fixations than less experienced musicians. Turning to more recent work, the findings of Goolsby (1994a), studying the sight-singing performance of 24 experienced musicians, equally divided into skilled and less skilled sight-singer groups, are again in agreement with Lehmann and MacArthur in terms of fixation duration, but in this case show no significant difference for fixation frequency. Truitt *et al.* (1997) researched experienced and less experienced pianists sight-reading single-line melodies and obtained the same broad pattern of results as Goolsby. In contrast to the other studies, Gilman and Underwood (2003), researching skilled and less skilled readers (controlled for technical skill at Grade 8 standard and above) performing dual-stave material on the piano, found no significant difference for fixation duration between their skill groups. Skilled readers did use 21% fewer fixations to carry out the task, though, but this needs to be interpreted in the context of a 22% faster performance speed, meaning that the overall rate of fixation was almost identical between the groups. The issue of how tempo influences eye-movement behaviour would seem to have been little considered by researchers, but is potentially crucial. Kinsler and Carpenter (1995) found that increasing tempo led to shorter fixation durations during the reading of rhythmic, non-melodic, notation. Souter (2001) provided evidence that both the number and duration of fixations made by skilled readers during the sight-reading of four-part hymn tunes decreased when tempo was increased. This evidence of a general dependency of eye-movement measures on tempo means that any comparison of behaviour based on skill groups performing at a single tempo risks being unrepresentative. For any clear understanding to be gained, investigation at a range of speeds would seem essential.

The range of results described in the last paragraph were obtained using a range of participants, methodologies, tasks, musical materials and performance tempi, making attempts to draw specific conclusions somewhat precarious. There would certainly seem to be no clear justification here for Lehmann and McArthur's assertion that skilled readers typically perceive larger groups of notes within their fixations. The findings of Goolsby and Truitt *et al.* are in fact entirely consistent with the idea that both better and less able sight-readers obtain similar levels of perception from their fixations, with less skilled readers simply taking longer to achieve this. Indeed, neither of these pieces of research found any significant difference between their participant groups in any of the other eye-movement indicators recorded except for the eye-hand span, a measure that will be discussed shortly. The same is true for the study by Gilman and Underwood. Furthermore, Truitt *et al.* recorded no significant variation in the pattern of fixation locations of the skill groups; that is, there was no clear evidence of any difference in information gathering strategy. Gilman and Underwood did not specifically study variation between their skill groups in this regard. Goolsby, interpreting the pattern of fixation location of his groups, found that 'skilled music readers look farther ahead in the notation, and then back to the point of performance, when sight-reading (1994a, p. 77). Evidence for this was far from conclusive, though, and to enable a clearer distinguishing between skilled and less skilled performance, in a subsequent paper (Goolsby, 1994b), Goolsby compared the eye-movement behaviour of his poorest participant from the former study with that of a top-performing participant. The skilled reader chosen was one of only two that had scored 111 marks out of 112 on the standardised sight-singing test used to categorise participants, and the less skilled reader had scored only 39 on the same test, clearly an extremely weak sight-singer. From his comparison of their data, Goolsby extrapolated that, 'skilled readers do not look at every note/rest in order to perform accurately. The less-skilled readers fixate on as much of the notation as time allows... progress note by note and perform with numerous errors (1994b, p. 120). The validity of such a generalisation would seem questionable, though, considering the possibly unrepresentative choice of participants.

Goolsby's work is widely cited in the literature, and would seem to be the principal reference for two ideas commonly presented in reviews: firstly, that less skilled readers generally fixate upon every note as they sight-read, and secondly, that the performance of skilled sight-readers is characterised by a greater strategic planning of eye movements than that of less skilled readers. From the other research considered so far, the former has already been shown not to be a typical finding amongst less skilled sight-readers. With regard to the latter idea, it is not only the unrepresentative choice of participant that is a concern; questions must also be raised about the degree of confidence that can be attached to Goolsby's interpretation of the eye-movement patterns. He, himself, admits the tentativeness of his interpretations, emphasising the fact that eye-movements themselves only provide clues relating to perception. He concludes that, 'the results of this study generate far more questions than answers regarding the music reading process' (Goolsby, 1994a, p. 94). Reviews of the literature have perhaps tended to be less circumspect in relation to his findings, though. There has also been a tendency for the variation in eye-movement behaviour observed by Goolsby to be interpreted as being purely perceptual in origin, when in reality the source may equally have been output related. For example, considering that his less skilled participant was in fact an experienced musician, it is quite possible that his fixating upon every note did not reflect a lack of input-related skill or strategy at all, but simply represented the need to look somewhere whilst attempts were made to effect the required pitches. As was discussed earlier, such a focus upon purely perception-based explanations for perceptual measures gathered from complete task performance has been typical of the interpretation of research generally within the domain, and is clearly an issue that needs to be addressed.

The final work that needs to be mentioned in relation to patterns of fixations is the perceptual sub-skill research conducted by Waters, Underwood and Findlay (1997) and Waters *et al.* (1998a), the former cited by Lehmann and McArthur (2002) as their principal reference for skilled readers using shorter and fewer fixations, and hence larger perceptual groups of notes, than less skilled readers. A close scrutiny of this study raises doubts about its ability to support these conclusions, though. Skilled and less skilled readers (controlled for technical skill) as well as non-musicians were required to visually compare monophonic musical sequences on a computer screen that were either identical or else differed by a single note, and to decide as quickly as possible whether the two were matching or not. The results fail to show any statistically significant difference between skilled and less skilled participants on the principal eye-movement measures, for example, number of fixations, fixation duration, and the number of times the participants 'flipped' their gaze between the two patterns on display. It is only when the two musician groups are compared with the third non-musician group, all of whom were previously unfamiliar with music notation, that significant differences on these measures are found. Overall, therefore, it is only the non-musicians that can be shown to have employed smaller groups of notes in their comparison strategy. Gilman and Underwood (2003) similarly found no significant difference in the recorded eye movement measures between skilled and less skilled readers for a perceptual task involving finding notational errors within dual stave material.

Waters *et al.* (1998a) carried out another pattern-matching experiment, this time without the use of eye-movement measuring technology, using dual stave stimulus and three groups of sight-readers, good, average and poor, who were reasonably controlled for technical skill. It was found that the good reader group outperformed the poor reader group (not the average group), requiring significantly less time to compare the two displays of notation. The researchers conclude this 'suggests that more skilled readers utilize larger musical units than less skilled readers, at least in this comparison task' (Waters *et al.*, 1998a, p. 137). However they do not really have *grounds for such a conclusion, because it is equally possible that the poorer readers in this experiment were simply using the same sized units as the other groups but* perceiving and processing at a slower rate.

An eye-movement measure that has been of particular interest to researchers has been the eye-hand span, which is the horizontal distance between the centre of the eye fixation and the note being simultaneously performed. Interpretation of findings specifically in relation to sight-reading skill is again made difficult because of the varieties of participants and methodologies employed. Jacobsen (1941) measured the eye-hand span of skilled and unskilled sight-singers to be up to 4 notes and 2 notes respectively. Weaver (1943) observed that on dual stave music the average eye-hand span was influenced by the complexity of the stimulus, but was on average 1 to 2 chords for the particular participants he investigated, with a maximum extension of 8 notes/chords. Goolsby (1994a) did not time-link eye-movements to performance and therefore no valid eye-hand span measure can be obtained from his sight-singing data. With monophonic material, Truitt *et al.* (1997) quantified the average eye-hand spans of their experienced pianist participants at 2 notes (their experimental materials consisted largely of crotchets), and that of their less experienced participants at 1 note. Furneaux and Land (1999) studying dual stave material, and comparing professional and novice/intermediate musicians, recorded eye-hand spans of 4 notes on average for the former and 2 notes for the latter. Gilman and Underwood (2003) recorded a mean eye-hand span of 1 beat for skilled readers and 3/4 beat for less-skilled readers with dual stave material. Some of Gilman and Underwood's beats consisted of a single crotchet chord, others a quaver chord together with a quaver passing note, meaning that it is not possible to translate their measures into an exact quantity of notation. These results indicate, overall, that the eyes of better readers fixate further ahead of the point of performance. However, with the research typically not controlled for technical skill, the extent to which this is relevant to explaining sight-reading ability cannot be gauged.

For Furneaux and Land, the larger eye-hand spans of experienced musicians are evidence of their ability to simultaneously process a greater quantity of notation from the score than less experienced musicians. They propose a short-term memory based account of sight-reading skill, suggesting that able readers have either larger storage buffers than less able readers, or else similar sized buffers and more efficient storage mechanisms. However, the work of Truitt *et al.* and Gilman and Underwood suggests the possibility that the larger eye-hand spans of able readers are simply indicative of their seeking information further ahead in the score, neither of their skill groups differing in terms of the quantity of notation effective to performance that they perceived within fixations (this aspect to their research will be discussed later). These researchers do not account for the different eye-hand spans of their able and less able readers, though, and clearly further research is required into this. One possible explanation, suggested by work into touch-typing, is that the size of the eye-hand span may be related to the degree of automation of motor execution. In touch-typing research, a measure called the stopping span, the number of keystrokes typed after a stop signal has been given, is considered reflective 'of the size of the execution buffer, which probably contains detailed parameters of movements that are no longer subject to much control or modification' (Salthouse, 1986, p. 310). The capacity of the stopping span has been found to be between one and two keystrokes (Logan, 1982), figures equivalent to Truitt *et al.'s* monophonic eye-hand span measures, and also to increase with typing skill (Legrand-Lestremau, Postal & Charles, 2006). The implication is that a larger execution buffer frees up working memory resources, enabling them to be directed at processing material further ahead of the point of performance.

To conclude this section, the interpretation of the eye-movement literature by the patterning account proponents, summarised earlier, would appear not to be an accurate reflection of research in this area. The range of studies carried out varies considerably in relation to the types of participants investigated and methodologies employed, meaning that there is, in reality, little in the way of clear insight that can be gained specifically into how skilled and less-skilled readers differ in their looking behaviour. The evidence available suggests the possibility that rather than having a fundamentally different pattern of response to notation, less-skilled readers instead may simply be slower at basic note processing from the score than skilled readers. Significant differences in the overall pattern of eye movements would seem to occur only when greater extremes of technical skill and musical experience are involved. Lee comments that eye-movement study in music reading remains in 'a state of gaze phenomenology without providing a viable theory' (Lee, 2004, p. 37). Given the sheer lack of research that has been carried out, together with its eclectic nature, it is hardly surprising that there is such a dearth of theoretical context within which empirical observation can be interpreted. Hopefully, as more studies are undertaken and lessons are learned from the methodological weaknesses of earlier work, a more robust knowledge base that is capable of sustaining theory development will emerge. However, to move beyond 'gaze phenomenology' it is also essential that research be carried out not merely to observe eye-movement behaviour, but to gain a more detailed understanding of the actual perception that readers are achieving from it. I will now turn my attention to a consideration of current knowledge in this area.

2.2.1.4 Measures of perceptual uptake

There have been three perceptual measures that have been the particular focus of research:

- The note identification span
- The eye-hand or eye-performance span
- The perceptual span

The note identification span

Bean (1938) measured the note identification span (what he termed the 'span of apprehension') of musicians, outside the context of normal sight-reading performance, by presenting them with displays of monophonic sequences lasting about 200ms (shorter than a typical eye fixation) and requiring them to subsequently perform on a piano keyboard as much as they could remember. He discovered that more experienced musicians had mean spans of approximately 5 notes (with a maximum figure of 9 notes) whereas less experienced musicians could recall only 1 to 2 notes on average. Concerning this experiment Sloboda wrote, 'Bean's study does not allow us to pinpoint the causes of this superiority as precisely as we would like. Does the superiority arise because experts have more rapid perceptual coding processes; or because they have better and more economical ways of storing what they perceive in memory; or because they have more efficient motor programs to organize the response? All the stages are plausible locations for the superiority. Subsequent research has helped clarify the situation' (Sloboda, 1984, p. 224).

The subsequent research referred to was a study by Sloboda (1978a) involving a development of Bean's basic methodology that attempted to provide answers to the above questions. Sloboda addressed the issue of motor influence on participants' responses by requiring these to be written down rather than performed on a keyboard. He also used a range of display times (20ms to 2s), which provided a broader context than Bean's methodology for interpreting the recorded measures. For example, Bean's single 200ms display may have only allowed time for a partial filling of short-term memory buffers, but Sloboda's approach provided the opportunity for a clear determining of asymptotic levels of storage. Unfortunately, the findings of Sloboda's research are not able to elucidate Bean's data as he claimed, because he employed different participant groups - musicians and non-musicians. Sloboda found that his musician group were more efficient at coding the notes from the display and could also store more notes in memory, the average maximum storage of musicians being 6 notes and that of non-musicians 3 notes. However, the meaning of Bean's results must continue to be ambiguous until further research similar to Sloboda's is carried out using truly equivalent groups of participants. In conclusion, therefore, there would seem to be no empirical justification for Thompson and Lehmann's statement that Bean's work shows that 'better and more experienced sight-readers remember longer sequences than less-skilled players' (Thompson & Lehmann, 2004, p. 147). At present, it cannot be ruled out that Bean's less experienced players were simply slower at encoding, the 200ms display providing them with insufficient time to demonstrate their actual storage capabilities.

The eye-hand or eye-performance span

Turning now to the measurement of perception gained in the context of the complete sight-reading task, I will first of all consider Sloboda's research into the eye-hand span (Sloboda, 1974). This is a different perceptual measure to the eye-hand span discussed in relation to eye-movement research, something that has led to confusion within other reviews and analyses of the literature. For example, Gabrielsson (2003) mixes and matches experimental findings relating to the two types of span, apparently assuming them to represent the same empirical measure. To prevent further confusion, I will from now on refer to Sloboda's species of eye-hand span by its less commonly used alternative name, the eye-performance span (Lehmann *et al.*, 2007). The eye-performance span is the quantity of notes that a musician can continue to perform if the score is unexpectedly removed, and may include information from more than one eye fixation. Proponents of the patterning account consider it to be an accurate estimate of effective preview i.e. the amount of preview essential to optimal performance (Sloboda, 1985). This accuracy cannot be guaranteed, though - as an estimate, it has the potential to both overstate and understate effective preview. In terms of the former, participants may perceive a greater quantity of notes than they actually need to organize their performances, something that may be more likely with lighter demand material. In terms of the latter, the memory trace available when the score is first removed may be subject to decay. Also, the span does not take into account partially identified notation, for example perceived within peripheral vision, that may nonetheless have a useful priming role in relation to subsequent processing and movement planning.

Sloboda's study involved performers on a range of musical instruments, informally controlled for technical skill, who were presented with previously unseen monophonic tonal melodies on a projector screen to sight-read, with reading ability level determined by the number of note errors in their performances. The projector was switched off at various points during their sight-reading, and participants were requested to continue playing as much as they could remember of any previewed material. Sloboda found that skilled readers could typically continue playing for about 7 notes, whereas less skilled readers were only able to continue for approximately 4 notes. These data were subject to a similar time constraint on perception as Bean's and Sloboda's note identification spans just discussed, one imposed indirectly by requiring participants to perform at a single, set tempo. Therefore, the same choice of explanation is possible as considered previously by Sloboda in relation to Bean's study: it could be that the skilled participants were either faster at encoding, had more efficient memory storage, or more effective motor responses. However, in contrast to his earlier careful analysis of Bean's results, Sloboda concludes the former two factors in combination to be the appropriate explanation for his eye-performance span data, in the apparent absence of any clear supporting evidence (Sloboda, 1984). There would seem to be no empirical justification for discounting a role for motor skill in explaining the variation in his measures. Furthermore, it cannot simply be assumed that the set tempo for the experiment allowed skilled and less skilled participants a sufficient time-window to fill short-term memory buffers, and so there is no guarantee that the recorded spans reflect their relative storage capacities; they may instead simply represent their relative rates of encoding. To be sure that the full extent of eye-performance span (and hence memory storage) achievable by participants is quantified, measurement at slower tempi is required until asymptotic values have been achieved. Researchers in this field, however, have tended to assume the eye-performance span does not vary with tempo. For example, Thompson and Lehmann write that, 'accomplished sight-readers consistently read around six or seven notes ahead (in a single-line melody)' (Thompson & Lehmann, 2004, p. 146).

Is there any evidence that the set tempo in Sloboda's study may have been too fast to allow eye-performance spans more representative of short-term memory storage capacity to be achieved? There is nothing relating to skilled readers that would especially indicate this, but Sloboda does document the fact that his less skilled participants found the required speed a very demanding one. He writes that they, unlike the skilled participants, 'often deviated markedly from the set tempo, either by going slower, or by making pauses where none were indicated in the text (Sloboda, 1974, p. 6). Whilst in the end no firm conclusions can be drawn about this particular issue without further research into the eye-performance span, it would certainly seem questionable whether such pressured performance conditions could have given a truly representative picture of these participants' perceptual and memory skills.

The perceptual span

Before discussing this work it is necessary to point out that the perceptual span is defined in different ways by different sight-reading researchers. For example, Lehmann *et al.* (2007) consider it to be the distance between the point of performance and the furthest point ahead in the score where the eye is obtaining information, and not necessarily confined to one fixation. Thus the perceptual span in this case will be either equal to or greater than the eye-performance span. However, some studies use the definition that is standard within text reading research: the quantity of notation effective to performance, perceived within a single fixation. This section focuses upon research into this latter version of the span by Truitt *et al.* (1997) and Gilman and Underwood (2003), no work apparently having been carried out into the former version.

Truitt *et al.* (1997) and Gilman and Underwood (2003) measured perceptual span using a 'gaze-contingent window paradigm' methodology, originally designed for research into text reading (McConkie & Rayner, 1975). In this methodology, participants' eye movements control the quantity of notation available to them on a computer screen. To the extent that they look ahead beyond the current note being performed, new windows of text, for example of 2, 4 or more beats in size, are revealed ahead of the point of eye fixation. Performances at these different window-constrained conditions are compared with a control condition in which notation is presented as in a normal music reading situation. The minimum window size at which performance and eye-movement measures equate to the control is considered indicative of the size of the perceptual span, defined in the typical manner of text research (Rayner, 1998) with the left boundary of the span as the point of fixation, the remainder extending to the right. Combining this span with the eye-hand span calculated from eye-movement data (discussed earlier) is considered to give the effective preview ahead of the point of performance, the measure for which Sloboda's eye-performance span is an estimate.

Truitt *et al.* (1997) studied monophonic sight-reading using perceptual windows of 2, 4 and 6 beats as well as an unlimited preview control condition. The experienced and less experienced pianist participants initially sight-read the control condition at an experimentally defined tempo, and were then required to perform the different window conditions as near to this speed as possible. Close to normal performance was achieved at the 4-beat window condition for both skill groups, pointing to perceptual spans of 3 or 4 beats beyond the point of eye-fixation. Combining these results with their eye-hand span measures (discussed earlier) reveals that their skilled readers were typically making use of the score 5 or 6 beats beyond the note they were actually performing, with less skilled readers making use of 4 or 5 beats (although the researchers themselves conclude that the lower figure in each case is more appropriate). Bearing in mind the different make-up of participants, probably little should be drawn from a comparison of these data and Sloboda's eye-performance spans. Gilman and Underwood (2003) employed similar technology to Truitt *et al.*, but in the context of dual stave notation (keyboard versions of 3-part chorales), using skilled and less skilled readers controlled for technical skill. They obtained perceptual spans for their two skill groups that when measured horizontally in beats were identical to Truitt *et al.'s* monophonic measures. Combining these figures with their eye-hand span data (also discussed earlier) makes the useful preview beyond the point of performance for skilled readers 4 or 5 beats, and for less skilled, 3 ¾ or 4 ¾ beats. As with Truitt *et al.*, the lower figure in each case is again considered more appropriate by the authors.

Gilman and Underwood were expecting a decrease in perceptual span from Truitt *et al.'s* figures because of the heavier task demand of the dual-stave material, something that would have been in line with findings in text research (Henderson and Ferreira, 1990). However, their perceptual span measure, although equal to that of Truitt *et al.* in terms of number of beats, suggested approximately a threefold increase in the quantity of notes processed, taking into account the multi-part nature of their material. For researchers using the gaze contingent window approach, perceptual span is typically assumed as marking the limits of short-term memory storage capacity. For example, two of the authors of Truitt *et al.* (1997), Rayner and Pollatsek, are senior researchers within the field of text reading, and in another paper relating to the same experiment they conclude that 'a major constraint on tasks that require translation of complex inputs into continuous motor 'transcription' is short-term memory. If the encoding process gets too far ahead of output, there is likely to be a loss of material that is stored in the queue' (Rayner & Pollatsek, 1997, p. 52). Such an assumption presents Gilman and Underwood with a conundrum: if Truitt *et al.'s* monophonic measures were definitive of memory constraints, how could it be that their own participants were able to exceed them by such a considerable margin? The solution they propose is that chunking mechanisms were invoked, participants achieving a lightening of perceptual and memory demands through sensitivity to familiar musical structure within the score. Whilst there may well be some truth in this proposal, Gilman and Underwood's conclusion that 'reading chordal harmony ... doesn't actually require more cognitive effort than reading a single-line melody' (Gilman & Underwood, 2003, p. 226) would seem somewhat overoptimistic.

A perhaps more appropriate solving of Gilman and Underwood's problem is achieved by questioning their underlying assumption that perceptual span is necessarily constrained by short-term memory size. Such an assumption does not in fact hold water on a closer inspection of their research. They seem to take for granted that all of the musical parts within their participants' 3 or 4-beat perceptual span 'windows' are perceived within a single fixation, but they have no empirical grounds for this. Work by Furneaux and Land (1999) has clearly established that in the reading of dual stave music, separate fixations are required for treble and bass staves, meaning that the memory representations driving the performances of Gilman and Underwood's participants must have been built up from at least two, and possibly more, smaller instances of perceptual span. Because of the multi-part nature of the material, the size of these fixations in terms of notes perceived cannot be established; all that can be safely concluded is their maximum horizontal extent, represented by the largest effective window size. In the absence of precise measures of individual spans, therefore, Gilman and Underwood's conclusion that 'there is little evidence to suggest that good and poor sight-readers have variable perceptual spans (Gilman & Underwood, 2003, p. 230) clearly cannot be supported from their results. Despite this ignorance over the details of perceptual span, though, there would seem no reason to doubt that from whatever pattern of fixations were employed, both skill groups achieved similar levels of effective preview overall, storing in memory information from a similar numbers of notes.

With the multi-part memory storage of Gilman and Underwood's participants made up from more than a single instance of perceptual span, it would follow that short-term memory capacity was probably not a factor limiting to the monophonic performances of Truitt *et al.*'s participants. These latter participants clearly needed no more than 3 or 4 notes of perceptual span to achieve normal performance levels, but it is likely that the availability of spare memory capacity offered the potential for further perception of a non-essential nature to be gained from individual, or indeed from additional fixations. The idea that some perception gained by sight-readers may be non-essential to performance output has scarcely been considered within the sight-reading research literature; the general impression given is that input is something constantly in short supply. This is not the case in other transcriptional domains though. In touch-typing research (Salthouse, 1986), for instance, the copy span, which is the quantity of text that can be typed from a single glance of the copy, which 'may correspond to the approximate working memory capacity of the typist' (Salthouse 1986, p. 309), is found to be consistently and significantly larger than the quantity of previewed text required to maintain maximal levels of performance.

Gilman and Underwood's results also question Sloboda's conclusion that short-term memory capacity was the limiting factor to the monophonic reading ability of his less skilled eye-performance span study participants. It was suggested earlier that Sloboda's single experimental tempo may not have allowed sufficient time for memory buffers to be filled, and that the smaller eye-performance spans of less skilled readers may therefore simply have been indicative of slower processing, not necessarily of smaller short-term storage capacity. Gilman and Underwood's results are consistent with this idea. The similar effective previews of the two skill groups, both larger than Sloboda's monophonic eye-performance span measures, were obtained in the context of skilled readers performing at a speed nearly one-third faster than less skilled readers, providing the latter with more time to process the score and therefore to give a more representative demonstration of their memory storage capabilities. Clearly, no firm conclusions can be drawn about this, but the evidence from Gilman and Underwood's study is also in line with evidence from eye-movement research pointing to skilled readers achieving faster input than less skilled readers, but in the context of a not dissimilar overall pattern of information gathering.

Truitt *et al's* and Gilman and Underwood's studies provide valuable insights into sight-reading skill, but it would seem that the research methodology and associated conceptual assumptions have been transferred from a text reading context perhaps without sufficient consideration given to the different nature of the music reading task, resulting, as has been seen, in considerable confusion in their attempts to interpret results. For example, in the context of the large perceptual spans characteristic of normal text reading, typically 14 letters (Rayner, 1998) it would seem quite appropriate for performance variation to be interpreted principally in terms of the contents of individual fixations. However, with Truitt *et al.'s* work indicating that considerably smaller perceptual spans are the norm with single line music reading, it would seem that whilst the size of perceptual span may still be relevant to explaining skill, emphasis also needs to be placed upon musicians' ability to piece together the contents of different fixations to provide a large enough memory representation from which successful performance can be planned.

In terms of the perceptual research as a whole, it is difficult to draw any clear conclusions about how skilled and less skilled readers differ in their skills. As with the eye-movement research, the number of studies is very small, and involves the use of different participant groups, materials, measures and methodologies that fail to provide anything approaching a coherent picture. The situation has caused Gabrielsson to comment on the "differences regarding concepts, choice of musical material, and measurement techniques which are in need of discussion before future research" (Gabrielsson, 2003, p. 244). The evidence is again sufficient, though, to raise questions about the proposal of the patterning account that skilled reading is necessarily dependent upon the achieving of larger quantities of preview, Gilman and Underwood's results suggesting the possibility that skilled readers are simply faster than less skilled readers at processing similar quantities of preview. It is vital that considerably more work is carried out into visual perception in sight-reading so that some clarification of these issues can be arrived at.

2.2.2 Musical structure and sight-reading skill

2.2.2.1 Introduction

With research into eye-movements and visual perception providing only mixed support for the patterning account's proposal that the ability of skilled readers is dependent upon their achieving larger quantities of preview than less skilled readers, I will now turn to an examination of the empirical evidence for the mechanism that is considered to drive this larger preview - sensitivity to musical structure within the score. I will first of all examine the empirical evidence for patterning perception and chunking playing a primary role in explaining skill difference, and then consider research into inference and prediction.

2.2.2.2 Chunking and pattern perception

Sloboda (1991) writes that his patterning account of music sight-reading was developed under the particular influence of the work Simon and Chase (1973) into chess perception. These researchers investigated the influence of playing expertise on the ability of chess players to remember board configurations. They found that experts could recall a greater number of pieces from meaningful configurations than novices, with this storage consisting of larger chunks of meaningfully related information. On random board configurations, however, no significant difference was found between expert and novice performance. Sloboda considered that such findings may hold the key to explaining the different abilities of skilled and less skilled music sight-readers, and he performed two pieces of research specifically to investigate this: his musical proof-reader error experiment (Sloboda, 1976b), to be considered in a later section, and his eye-performance span studies (Sloboda, 1974, 1977).

The first of the eye-performance span studies (Sloboda, 1974) has already been discussed, in part. Sloboda proposed that what he interpreted as the larger memory storage of skilled readers in this study (on average, eye-performance spans of 7 notes) was due to efficiencies resulting from these readers being more sensitive to musical structure within the experimental materials than less skilled readers (on average, eye-performance spans of 4 notes). 'We may hypothesise', he wrote, 'that the poor sight-reader is unaware of, or unable to use, structures or redundancies in the text. His capacity is thus limited to 4-6 bits (five items) which Klemmer and Frick (1953) have shown to be the capacity for retention of dots in a two-dimensional matrix (which is what musical notation is if considered as a purely visual stimulus). On the other hand, the good sight-reader makes use of redundancies or structures to increase his capacity' (Sloboda, 1974, p. 6). The validity of such a proposal clearly rests upon Sloboda's somewhat unwarranted assumptions, discussed earlier, that his eye-performance span data are indeed representative of his participants' memory storage capabilities, and also that they are not significantly subject to influence by output mechanisms. Also, because he did not run a control condition using unstructured materials, he cannot be sure that it was sensitivity to musical structure that was responsible for his skilled participants' larger eye-performance spans.

Despite his methodology not involving any structural manipulation, Sloboda nonetheless found evidence of what he considered to be differential structural sensitivity between his skill groups in the data of this experiment. He writes that, 'the experimental design was such that participants were deprived of the score at various distances prior to a musical phrase boundary. It was found that there was a greater than chance likelihood of EHS coinciding with a phrase boundary. This effect interacted with reading ability... for good readers, the EHS was not constant; it expanded and contracted to accommodate a phrase unit' (Sloboda, 1984, p. 231). There are grounds for questioning the extent of this interaction, though. Firstly, consideration needs to be given to the considerably greater task load experienced by

Sloboda's less skilled readers in this study, discussed earlier, that would seem to make this study not an entirely fair test of their perception of phrasing. Also, in a more recent publication, Sloboda appears to contradict his earlier assertion. He writes that 'when the phrase boundary was at a reachable distance (say 6 notes), even those participants whose average range of eye-hand span was only three to four notes reached the boundary' (Lehmann *et al.*, 2007, p. 117). Such an increase represents a doubling of eye-performance span for his poorest sight-readers, a relative gain that is at least equivalent, if not greater than that achieved by his skilled readers (even his most skilled readers could not extend their eye-performance span to the phrase boundary if it was more than 10 notes away (Sloboda, 1974)). This quote somewhat undermines the case for the patterning account presented in the chapter from which it is taken.

In the second eye-performance span study (Sloboda, 1977), Sloboda measured the eye-performance spans of skilled readers performing not only tonally coherent music as before, but also structurally disrupted material that broke the rule of tonal progression, approximating to an unstructured control condition for this reader type. With this latter category of material, the skilled participants recorded smaller spans than with the former type, leading Sloboda to conclude that during its performance 'preview was not so useful, and cannot, indeed, be sustained at normal levels' (Sloboda, 1985, p. 72). Whilst the latter point is supported by the data, the former would appear more questionable. If preview was not so useful with the tonally disrupted material, the quality of participants' performances should have become degraded. However, this does not appear to have been the case. For example, Sloboda mentions a participant who 'had not experienced the atonal melodies as any more difficult to read than the others' (Sloboda, 1985, p. 197). If the disrupted melodies were no more difficult to read, though, there would seem no justification for regarding the performance of the tonal melodies as dependent upon their larger associated eye-performance span values; on the contrary, the extra preview recorded would appear to be surplus to requirements. So here is evidence even from Sloboda's own work that is questioning of the importance of larger preview to explaining sight-reading skill.

The findings in relation to this particular participant stand in stark contrast to Lehmann *et al.'s* own anecdotal experience of the difficulty of sight-reading unconventional material. They write, 'we know from our own experience that memorizing or sight-reading unconventional (e.g. non tonal) material can be extremely frustrating because memory skills are so specific. This effect is due to the breakdown of our chunking mechanisms, and instead of coding larger, meaningful units (tonal melodies and harmonies), we have to group individual notes or intervals (Lehmann *et al.*, 2007, p. 113). Sloboda does admit the possibility that the 'real time' demands of monophonic reading may be less than that of dual-stave, and therefore that the perception of structure may be 'less necessary' (Sloboda 1985, p. 78). If this is the case, it might suggest that Sloboda's atonal study participant was able to perceive sufficient structure from the disrupted sequences for his needs. However, it is also possible that the memory and processing speed demands of skilled monophonic reading are simply not sufficient to require structural perception and that the task is therefore largely achievable using basic information processing mechanisms.

Although there are questions regarding the extent to which Sloboda's skilled eye-performance spans are effective to performance, the data do indicate that skilled readers are sensitive to musical structure. However, in the absence of an equivalent atonal control condition for less skilled readers, there is no empirical basis for his conclusion that the smaller tonal eye-performance spans of these participants compared to skilled readers necessarily points to their being less sensitive to musical structure. Just because 'poor readers seem to behave with 'normal' music rather like good readers with 'obscure' music' (Sloboda, 1985, p. 72), it cannot simply be assumed that the similar outcomes have a common cause. It could equally be the case that less skilled readers were as sensitive to patterning as skilled readers but simply slower in their basic rate of encoding. The tonal eye-performance span study was Sloboda's only research involving less skilled readers who were able instrumentalists, and therefore it is clear that he developed his patterning account almost in the complete absence of evidence relating to the structural sensitivities of this type of reader.

There appears to have been only one other published study that has investigated the perception of musical structure amongst proficient musicians, and it fails to provide any support for the idea that less skilled readers are any less sensitive to musical patterning than skilled readers. As part of one of their pattern matching studies discussed earlier, Waters *et al.* (1997) manipulated the structure of their experimental materials in a number of ways to see what effect this would have on the speed and pattern of perceptual activity of their three skill groups - skilled and less skilled readers controlled for technical skill, and beginning instrumentalists. Comparisons involved identifying errors within monophonic tonally structured, rhythmic musical sequences (described as pitch coherent and temporally coherent) and within randomised versions of these sequences. The temporally randomised materials retained the same order of note pitches as in the coherent versions, but randomised the durations throughout each sequence. Pitch randomised materials contained exactly the same notes as pitch coherent materials, but in a randomised order, disrupting the previously coherent tonal structure. Regarding the temporal transformations, both skilled and less skilled sight-readers displayed sensitivity to the structural distinction, with only beginning instrumentalists failing to do so: "both expert groups were slower to respond to temporally randomised material...whereas the novices showed no such sensitivity to structure.' (Waters *et al.*, 1997, p. 481). With pitch transformations, all the skill groups were slower with randomised than with coherent materials but 'there was no evidence of an expertise x pitch structure interaction' (Waters *et al.*, 1997, p. 481). Lehmann *et al.* (2007) misrepresent the findings of this study, surprisingly turning them instead into unambiguous support for the patterning account: "not only were skilled sight readers faster compared with less skilled sight readers, but they were also more sensitive to disturbances from randomisations of tonal and rhythmic parameters. This handicapping effect of expertise underscores how strongly experts rely on the patterned nature of the stimulus' (Lehmann *et al.*, 2007, p. 117).

Whilst there is evidence, then, that skilled readers are sensitive to musical patterning, there would seem to be none available demonstrating that less skilled readers are any less so. Research would only seem to have found variation in sensitivity to structure amongst participants when skilled musicians are compared with novices, as in the case of Waters *et al.* (1997) quoted in the previous paragraph, or when musicians are compared with non-musicians (Halpern & Bower, 1982). In Halpern and Bower's study, participants were asked to memorise traditionally tonal and rather random melodies, with musicians only achieving better recall on the former type of material. Such results are clearly in line with Simon and Chase's chess study findings that originally inspired the patterning account, and in that they involve the comparison of experienced musicians and novices, would seem to be the true musical analogue of this research. Sloboda, and the other patterning account proponents, appear to have appropriated Simon and Chase's model of skilled memory to attempt to explain a skill difference to which it would not really seem appropriate. Less skilled readers who are otherwise expert musicians can in no way be considered true novices at the task. Although they lack ability, their sight-reading performance is necessarily carried out in the context of considerable musical knowledge and expertise. In the light of this, it is difficult to see how their lack of skill at the task has been so readily attributed to a lack of sensitivity to musical structure. Clearly, though, with the dearth of research specifically into musical patterning, no definitive conclusion can yet be drawn either way about its role in explaining the different abilities of skilled and less skilled readers. However, it is of no small concern that the patterning account has achieved such hegemony within the domain apparently in the complete absence of data indicative of less skilled readers actually being handicapped in this regard.

2.2.2.3 Inference and prediction

In his musical proof-reader error study (Sloboda, 1976b), Sloboda demonstrated that by slightly altering occasional notes in the score to be out of style within pieces of dual stave, tonally structured classical repertoire, such alterations would often be ignored by skilled reader participants, and replaced by notes more in keeping with the melodic/harmonic context. This was carried out typically without the conscious awareness of the performer. This experiment unambiguously demonstrates that skilled readers were not simply decoding the stimulus material note-by-note from the score, but were using prior musical knowledge and expectancies to reconstruct the music from their structural perception. However, the lack of any unstructured control in this study means that no conclusions can be drawn about the specific performance gains attributable to inferential mechanisms. In the context of this work, it is also worth considering a qualitative study of the introspections of four expert sight-readers by Wolf (1976). Wolf concluded that their skill was fundamentally dependent on the ability to both search out familiar patterns in the score and to make hypotheses concerning them. Both these pieces of research provide fascinating insights into how the perception of structure impinges upon skilled reading. However, no explanation for skill difference can be gleaned from them because no comparisons with less skilled readers were undertaken. The impression is sometimes given in discussion of these studies that experts' strong perception of structure is somehow evidence of less skilled readers' inferior perception. This is obviously a logical fallacy, though. We simply have no idea of the extent to which less skilled readers make proof-reader errors, or the manner in which they mentally represent the score.

Lehmann and Ericsson (1996) also researched higher-order structural influences, investigating memory and improvisational skills amongst 16 college level pianists, specialising either in solo performance or accompanying. To begin with, the sight-reading ability of participants was measured by requiring them to sight-read a number of pieces involving instrumental accompaniment. To investigate memory (recall), participants were asked to perform a piece that they had already played through a number of times during the study, but which on this occasion contained some sections of the notation left blank during which they were required to play what they could remember from previous trials. To study improvisation, participants were required to play a new piece similar in style to that used on the recall task but which included some blanked out sections, which they had to improvise their way through. The results for the recall task were 'consistent with scores on the sight-reading task' (Lehmann & McArthur, 2002, p. 141) and those for the improvisation task were significantly correlated with sight-reading ability. The findings for the recall task were considered evidence of skilled readers being able to learn material more quickly than less skilled readers owing to 'their superior ability to grasp the structure' (Lehmann & McArthur, 2002, p. 141). On the improvisation task it was considered that better readers again 'assimilated the structure of a piece with its redundancies' (Lehmann & McArthur, 2002, p. 142). There is no evidence within the research, though, that necessarily points to such structure-based explanations of Lehmann and Ericsson's results. It could equally be the case that difficulties with basic perceptual and motor elements of the tasks could have hindered the performance of the less skilled readers. Another experiment in the same piece of research, considered in more detail later, provides some support for this, less skilled readers being found to perform sight-reading related motor activity less accurately than skilled readers. In view of this, any test involving such motor output would likely have led to inferior performance by the former group whether or not use was made of higher-order structuring.

2.2.2.4 An assessment of the patterning account

So to what extent has research been found to support the following fundamental contentions of the patterning account?

- The perception of structure is essential for skilled levels of reading
- The weaker ability of less skilled readers is primarily the result of their being less sensitive to musical structure than skilled readers
- The perception of structure enables skilled readers to achieve both larger preview and to process it at a faster rate

Sloboda's dual stave musical proof-reader error study has clearly demonstrated that the perception of musical structure plays a role in skilled performance. However, the absence of a control condition using unstructured material of similar complexity means that we cannot be absolutely certain that the perception of structure demonstrated was essential to the level of performance achieved. In relation to monophonic reading, Sloboda's eye-performance span study indicates that skilled readers are perceptually sensitive to musical structure, but there is insufficient evidence to conclude that skilled monophonic performance is actually dependent upon it. It is possible, therefore, that rather than being a ubiquitous requirement for skilled output, structural sensitivity may instead become more relevant to performance as notational complexity, and thus task demand, increases. Turning to less-skilled readers, there would appear to be no evidence to substantiate the idea that less skilled perception or performance is in any way less sensitive to musical structure than that of skilled readers. Empirical work has only succeeded in demonstrating that musicians as an entire group are more sensitive to structure than beginning instrumentalists and non-musicians.

Finally, considering preview size, evidence has been presented suggesting that Sloboda's eye-performance span data may have been inappropriate as the principal empirical foundation for a theory of sight-reading, the methodology possibly having led to an understating of the extent of perception/memory storage achievable by less skilled readers, and also an overstating of skilled perception that is effective to performance. The work of Gilman and Underwood (2003) has indicated that the difference in preview use between skilled and less skilled readers may only be very small - one quarter of a beat in their research. If Gilman and Underwood's results are valid, it may still be possible that sensitivity to musical structure plays the primary role in accounting for the superior performance of skilled readers, but it would require such sensitivity to mainly lead to faster perception, rather than significantly greater perceptual capacity. But considering the lack of any clear evidence indicating less skilled readers to be less sensitive to structure than skilled readers, it is equally possible that skill difference may simply be attributable to basic perceptuo-motor factors and cognitive strategies unrelated to specifically musical content, with sensitivity to structure being only a secondary influence upon performance.

In conclusion, the currently available empirical data relating to visual perception and processing would seem far too ambiguous to provide a basis for any confident theorising about the roots of sight-reading skill difference. This begs the question as to how the patterning account could have been so confidently espoused by its proponents and have achieved such widespread hegemony within the domain. Perhaps the ideological context within which the original research was carried out may be relevant in trying to understand this. Research into music psychology prospered in the 1970s and early 1980s within a strongly structuralist cognitive paradigm, with researchers clearly excited about the explanatory power of such an approach for the study of music performance. For example, musical expression was understood principally in terms of generative processes driven by the perception of structural elements within the score (Clarke, 1985; Shaffer, 1984). It would seem that the mindset of researchers at the time was such that it was scarcely conceivable that non-structural factors would be of any significant relevance to the issues being investigated. Over time, however, new empirical evidence began to emerge in some areas that challenged prior assumptions, necessitating interpretation involving a wider range of influences. In relation to the study of musical expression, Clarke writes that earlier research 'had fallen into the trap of focusing too exclusively on the relationship between structure and expression' and that an 'increasing recognition of a multi-dimensional perspective' was required (Clarke, 1995, p. 53). Research into the psychology of performance expression has subsequently flourished within this broader context.

Given these significant developments in related areas of study, it might perhaps have been appropriate for patterning account proponents to consider whether sight-reading research had not fallen into a trap similar to that described by Clarke. However, with Sloboda's empirical focus upon sight-reading largely ending with research published in 1978, and little new work carried out in the ensuing fifteen to twenty years, there would seem to have been no empirical catalyst to provoke any questioning of the status quo. With more recent research, like that of Truitt *et al.* and Gilman and Underwood, this is no longer the case, though, and it is some cause for concern that in his most recent presentation of the patterning account (Lehmann *et al.*, 2007), Sloboda has not engaged with the findings of this work at all. Clearly, it does not follow from this recent evidence that the patterning account is necessarily wrong; as has been said earlier, there is simply too little research available for any firm conclusions to be drawn. But equally there is insufficient evidence that it is correct. It would seem vital, therefore, that research is carried out that will enable an adequate testing of the account. If it is confirmed by such research, this will be all well and good, because it will then be provided with the empirical backing that it currently lacks. If, however, musical structure is found to be less relevant to explaining skill difference, then the ensuing 'multi-dimensional perspective' would liberate research from its currently rather narrow confines to begin examining a broader variety of perceptual, memory processing and motor mechanisms which may potentially play central roles in explaining skill difference, but that up to now have been considered of only secondary importance.

2.2.3 Auditory Representations

Sloboda (1978a) studied musicians and non-musicians on an interference task involving the memorising of sequences of letters and musical tones whilst simultaneously writing down a briefly displayed segment of musical notation. The former group were significantly better at the task, and Sloboda concluded that this superiority arose from their use of a 'non-verbal, non-acoustic type of memory' (Sloboda, 1978a, p. 14) for musical notation. This led him to propose that skilled sight-reading typically was not dependent upon the use of acoustic representations. However, other studies (Lee, 2004; Waters *et al.*,1998a; Kornicke, 1995) have indicated the use of auditory representations within skilled reading, and claim a relationship between the ability to form auditory representations from musical notation, and sight-reading skill. Lee (2004) found that this ability explained 18% of the variance amongst her participants' sight-reading scores and in Kornicke's research, it was the highest predictor variable, responsible for 15% of total variance. Waters *et al.* speak for all these researchers when they conclude that the correlations obtained show that 'generation of auditory representations from visual structures in the score has some role in skilled reading' (Waters *et al.*, 1998a, p. 143), implying that it plays less of a role amongst less skilled readers. Such a conclusion is not securely supported by any of the experimental evidence, though. Forming an auditory representation from notation, by definition, has dependencies upon skills relating to the initial visual input of that notation, and so results from such a task are necessarily biased against less skilled readers who may have difficulties with such input. Further research is needed in which auditory imaging is studied in greater isolation from this potential source of variation.

Two possible roles for auditory representations within sight-reading have been considered by researchers, one relating to performance feedback, the other to possible involvement of auditory representations in the actual directing of motor output. Wolf proposed the former, writing that, 'hearing the music appears to be a kind of verification mechanism. It allows the musician to make sure that the transfer from eyes to fingers has gone smoothly and accurately' (Wolf, 1976, p. 154). Banton (1995) has found experimental evidence for this. Comparing normal piano sight-reading performance with a condition in which no auditory feedback was available, there was no evidence of a significantly greater number of execution errors on the latter. However, there were signs that during normal performance better sight-readers were able to detect by aural means the beginnings of deviation from the score, which enabled appropriate correctional motor adjustments to be made.

The idea that auditory representations derived from notation may be used to drive motor output has been proposed by Kornicke (1995). Although her research, as discussed above, does not specifically lend support to such a process, she considered that there was a logic to the idea. She writes, 'music involves the conversion of printed notation into sound. It would appear that individuals who could more easily form a mental image of the sound from printed notation would have an advantage in sight reading musical scores' (Kornicke, 1995, p. 72). In support of Kornicke's contention, one of the four expert sight-readers interviewed by Wolf (1976) stated that his reading skill was principally dependent upon auditory representations directing motor output. 'If you know what the piece should sound like...then you know the patterns that your hands should be playing (p. 159). However, empirical support for this mechanism awaits future study.

2.3 Output mechanisms: psychomotor skill

The idea that sight-reading skill difference may be caused by variation in motor ability has scarcely been considered in the published literature. The dominating cognitivist mindset would seem to have caused researchers to assume that the skilled motor programming visible in rehearsed performance would necessarily follow through into the production of movements at short notice in the performance of sight-read material. Waters *et al.* make this point explicitly: '... the fact that musicians can have similar general performance abilities ('output' skills) but vastly different sight-reading abilities ('input + output' skills) implies that the attainment of input skills must be important to sight-reading facility' (Waters *et al.*, 1998a, p. 125). In consequence, research has focused almost exclusively upon perceptual and cognitive factors. One might conjecture that the use of the term 'sight-reading', with its implication of a purely mental appropriation of meaning, has had a subliminal constraining influence upon researchers in terms of both the type of study that has been undertaken and their analysis. Perhaps the more holistic 'performing at sight' or 'sight-playing' would be a more appropriate and representative designation of the task.

There seems to have been very little research into the motor ability of skilled and less skilled sight-readers. Sloboda, Clarke, Parncutt and Raekallio (1998) studied the fingering strategies of pianists of different levels of expertise. Whilst this work provides a fascinating insight into issues relating to motor planning and ability, it is of limited value in gaining a foundational understanding of the sub-skill because participants sight-read from notation, with responses therefore subject to input skill variation. Lehmann and Ericsson (1996) claim to have studied motor skill in a more isolated manner, though. They required their college level pianists, discussed earlier, to perform a 'leap task', a study of 'spur of the moment' kinesthetic ability, tested by measuring the accuracy of participants' performances of partially rehearsed sight-read extracts involving jumps across the keyboard. This was undertaken both with and without visual feedback of hands and keyboard, the latter condition enabled by participants wearing adapted goggles that limited the field of vision to the presented

UNIVERSITY
OF SHEFFIELD
LIBRARY

notation only. The results of both conditions were strongly correlated with sight-reading ability, the researchers concluding skilled sight-reading to be less dependent upon visual monitoring than less skilled reading (something that they considered to be acquired during instrumental training, rather than specifically though sight-reading training). Both skill groups performed less well on the leap task involving no visual feedback, indicating that some vision of the hands and keyboard to be necessary to achieve optimal performance. Banton's study (1995), part of which was discussed in the previous section, also demonstrated this. Her work involved a further 'blind' condition where participants were unable to view the keyboard during playing, leading to a greater number of wrong notes performed.

Lehmann and Ericsson's conclusion that skilled readers are less dependent upon visual monitoring than less skilled readers is not the only possible interpretation of their empirical evidence. First of all, despite their claim, motor activity was not sufficiently isolated from notational input in their study, and therefore participants' output could have also been subject to variation as a result of differential skill in the initial perception and processing of the notation. Secondly, it is also possible that some of the performance difference may have resulted from the playing of skilled readers being less dependent upon rehearsal - something not discussed in their paper. Since this piece of research, Lehmann appears to have developed his thinking in this area, because in a recent chapter about sight-reading (Thompson & Lehmann, 2004), the authors emphasise the difference between the online, unrehearsed motor program production required for sight-reading and keyboard improvisation, and that employed in rehearsed performance. However, although they consider that ability at the former may be a factor in reading skill determination, it would appear that the use of more automatic motor sequences, based upon long-term memory structures, is still viewed as fundamental to the task. Given the lack of emphasis upon output-related research within the domain, it would seem particularly important that further empirical studies are carried out in this area. For example, it is necessary for research to gain clearer insight into the extent to which skilled and less skilled readers differ in their dependency upon visual monitoring by achieving a greater isolation of motor output from potential variation due to input factors. Also, work is needed to empirically test the idea that ability at unrehearsed movement is a relevant factor to reading skill variation.

As mentioned earlier in the chapter, though, the apparent openness of these authors to motor-driven variation in sight-reading skill is potentially undermining of their central thesis concerning the primacy of the perception of musical patterning in explaining reading skill difference. If they admit motor ability to be an important source of skill variation, they cannot logically continue to interpret the perceptual measures upon which the patterning account is based, for example, the eye-performance span, in a manner that ignores the possible influence of motor factors in their determination. A willingness to entertain a role for motor influences in determining sight-reading skill necessitates, therefore, some review of their interpretation of these perceptual measures.

Like motor skill, the study of kinesthetic and motor imagery has similarly been little considered by researchers. Another of Wolf's expert sight-readers claimed that his skill depended particularly upon the use of kinesthetic representations. '...we have a kinesthetic imagery. That means you feel ... the positions of the black and white keys, the stretches of octaves and other intervals, scored positions and things of that kind [and they] have a very precise mental image' (Wolf 1976, p. 159). Such representations, however, have not sat comfortably with the cognitivist paradigm that within which much of the research into sight-reading has been undertaken. With the growth of interest in embodied imagery within psychology in general (for example, Jeannerod (1994) and Glenberg (1997)), the study of musical sight-reading would seem to offer a potentially fruitful area of focus.

2.4 Research within the expertise paradigm

Within sight-reading research focusing on factors predictive of expertise, there are three main pieces of work to be considered, those of Kornicke (1995), Lehmann and Ericsson (1996) and Lee (2004). Aspects of these studies have already been considered - the principal emphasis of this section is on findings relating to the role of practice in skill determination gained from biographical evidence. Broader elements of Lee's work are discussed, though, which have not so far been relevant to discussion.

Kornicke (1995) found that sight-reading skill variation was related to the amount of sight-reading experience of her participants. Investigating this in more detail, Lehmann and Ericsson (1996) found that it was specific types of experience that were of particular importance in skill determination: accumulated accompanying experience and size of accompanying repertoire, which combined accounted for 61% of the variance in their participants' sight-reading scores. The authors argue that the relationship is a causal one. 'Individual differences in sight-reading ability in our participants and exceptional sight-reading feats by eminent musicians do not seem to reflect innate music talent or a specific sight-reading talent. Rather they are the results of deliberate long-term involvement in relevant domain-related activities and appropriate self-imposed challenges' (1996, p. 25). In other words, skilled readers owe their ability primarily to the manner in which they have developed their skills through practice and collaborative performance. Although Lehmann and Ericsson claim to have found evidence of other related sub-skills that are linked to sight-reading variation, for example, sensitivity to higher order structuring discussed earlier, differential ability at such sub-skills is itself regarded as practice driven. This research is very useful in demonstrating that sight-reading cannot be considered merely a natural talent that skilled readers develop effortlessly; becoming an expert clearly involves considerable work and commitment. However, the authors' rather sweeping conclusions do go somewhat beyond what is actually justified by their research data. They have certainly shown that practice is essential for skill development, but there is insufficient evidence to conclude that quantity of appropriate practice is the fundamental cause of skill and that individual differences are not important to the explanation. For example, they do not give sufficient consideration to the reasons why there is differential engagement in practice. It is possible, for example, that individuals who are particularly disposed to practising sight-reading are those who experience a greater ease at the performance of the task, and sense that their practice is leading to significant skill improvement. By contrast, lack of practice at sight-reading may result from individuals finding the task problematic in some way, and therefore deciding to focus their efforts upon developing alternative, more readily achievable, musical skills. For any confident conclusions to be drawn about these issues, it will be necessary for long-term longitudinal studies into sight-reading skill development to be undertaken.

Lee (2004) studied a wide range of range of factors potentially influential upon skill (25 in all). As well as recording relevant biographical practice data for her participants, who were all university music department based, she undertook a variety of tests to measure aspects of their perceptual, cognitive and psychomotor functioning. The three factors that were found to explain the largest proportion of the variance in sight-reading scores were the speed of trilling achieved with fingers 3-4-2 (33% of variance in sight-reading performance), the speed of trilling with fingers 3-4-1 (26% of variance), and quantity of deliberate practice at sight-reading up to the age of 15 years (25%). All the other factors explained proportions of the variance of less than 20%. The age specific finding in relation to practice suggests 'the importance of sight-reading expertise at primary and secondary school levels' (Lee, 2004, p. 137) i.e. without secure foundations laid early, older musicians may be limited in the progress they can make at the task. Lee presents a number of explanations as to why the 'speed trills' should be so significant in explaining sight-reading skill variation, for example instrumental performance expertise transferring over to sight-reading achievement and an individual's tremor speed (fixed from birth). However, it would also seem possible that these results simply reflect a lack of adequate experimental control upon the performance ability of her participants, their technical skill range being spread across a spectrum from first-year students to staff. On average, one would expect that those further advanced in their course of study would be both better performers - thus faster trillers - and better sight-readers.

Concerning cognitive factors, both music-specific and non-music specific short-term memory provide only an insignificant proportion of overall variance, with working memory making up only 6%. This result provides some confirmation of my own questioning of sight-reading accounts based on short-term memory capacity.
A number connection test, which is closely correlated with general intelligence, accounted for 19% of variance. Overall, as well as demonstrating again the importance of practice to skill development, Lee has provided evidence that sight-reading ability has possible dependencies across a range of component skills, and the study is extremely valuable in highlighting potentially relevant factors for future research to focus upon. However, as with Lehmann and Ericsson's research, conclusions regarding a fundamentally causal role for the significant factors cannot be drawn from mere evidence of correlation.

2.5 The measurement of sight-reading ability

To be able to confidently conclude that variation in a particular performance indicator is associated with variation in sight-reading ability, one needs to be sure that the measures of sight-reading ability employed are valid ones. This section examines how sight-reading ability has typically been quantified in previous work, some of the problems associated with the methods used, and some indication of how these might be addressed in future study.

The typical empirical method used to measure skill has been to require participants to play a previously unseen piece of music, and then to mark the performance on a single linear scale. Sloboda (1974) based his quantification of skill upon a single factor - the number of pitch errors. Other researchers, for example, Waters *et al.* (1998a) have attempted a more comprehensive representation of skill by making separate assessments of the accuracy of pitch, rhythm and of musical expression, and then combining them to provide a single score for each participant. Using such a single-dimensioned scale might well be appropriate for the most skilled sight-readers, for whom a very high overall mark will necessarily be indicative of expertise across all components. However, lower overall marks achieved by less highly skilled participants could represent a variety of component combinations, meaning that any distribution of marks would be limited in terms of the detailed information it could convey about skill variation. For example, some less-skilled readers may be particularly weak at deciphering rhythms (Elliott, 1982), whilst others may find pitch determination more of a problem. It is also possible that pitch finding itself may not be a single skill - individuals might vary in ability with different types of stimulus, for example, monophonic, contrapuntal and chordal. Having assessed and marked the separate components, there is then the difficulty of deciding upon the weighting they should be given within the overall mark scheme. For example, Waters *et al.* devoted 50 per cent of the overall total to the measurement of expression, a figure that would seem rather excessive. Bearing in mind that it will tend to be skilled sight-readers who perform more expressively because of their greater level of basic fluency at the task, such a weighting would likely have led to a considerable skewing of the mark distribution in favour of this type of reader, something that may question the validity of its use within subsequent correlation based analysis.

It would appear, therefore, that meaningful measurement of sight-reading necessitates some form of multidimensional scale. However, this still does not quite provide an acceptable approach because the errors present within a performance may not always be symptomatic of participants' particular skill deficiencies. There has been little research into this, but evidence suggests that, for example, pitch errors may result from interference effects due to difficulty in deciphering rhythms (Waters, Townsend & Underwood, 1998b). Indeed, to cope with more challenging stimulus material, skilled readers may intelligently weave improvisation into their performances (Sloboda, 1976b), resulting in what technically are errors, but of a strategic and musically meaningful kind. It would seem, therefore, that a valid measurement of sight-reading skill would also involve some assessment of rhythmic and pitch components in isolation, together with a means of judging whether pitch errors have contextual meaning, or are simply mistakes.

The complexity and required performance speed of the materials used also need to be considered carefully. If they are too facile or the set tempo too slow, this will likely have a ceiling effect upon the overall distribution of marks awarded, overstating the abilities of less skilled participants. If the materials are too difficult or the set tempo too fast, this may result in floor effects, particularly for less skilled readers. Indeed, with material that is too challenging, the performances of the latter group may be considerably degraded, perhaps leading to a substantial understatement of actual skill level. A possible example of this has already been mentioned in relation to the eye-performance span work of Sloboda (1974), where his less skilled participants struggled to sight-read at the same tempo as his skilled participants. Choosing stimulus material and a tempo equally suited to measuring the ability of both skilled and less-skilled readers therefore requires careful consideration. Lee (2004) has attempted to address this particular issue by measuring her participants on a selection of reading tests across a range of difficulty levels.

Because of methodological weaknesses relating to sight-reading ability measurement, the results of many of the studies that have been discussed in this review need to be interpreted with caution. It would seem likely that much of what research has revealed about skilled readers is reliable; however, knowledge concerning less skilled readers may be more questionable, with possible implications for any theoretical understanding of sight-reading variation that has been built upon the available empirical data. This issue will be returned to in the next chapter where an alternative method of sight-reading measurement is proposed that would appear to offer some resolution to a number of the problems that have been raised here.

2.6 General conclusion and areas needing research

This review of the literature has shown that the history of sight-reading research has not been one of stepwise development, in which a consistent and secure body of knowledge has been built up. Empirical work has instead been somewhat piecemeal in nature, with researchers approaching the subject from different angles, with different ideologies and methodologies. Studies have also had a tendency to investigate more arcane aspects of the task before more basic, foundational issues have been properly understood. Problems are further compounded by the methodological and conceptual weaknesses of some studies making their evidence and conclusions possibly unreliable. A final point to mention is simply the lack of studies available, which compromises not only the extent of knowledge, but with the lack of sufficient replication of research findings, also its security. As a result of these factors, it is difficult to gather the strands of research evidence together into a meaningful whole for an explanation of why skilled sight-readers are better at the task than less skilled readers. There are certainly plenty of clues available, but still far too many gaps in knowledge to make the development of theory anything more than a tentative, hypothetical affair. This review has demonstrated that the drawing of definitive conclusions about the origins of sight-reading skill must necessarily await a considerably more exhaustive investigation of basic perceptual, cognitive and motor related measures. Such a resource has been collected within the related transcriptional domain of touch-typing (Salthouse, 1986), enabling theoretical understanding of that particular task to progress far in advance of what has been achievable in music reading.

In the light of this, the focus for this current research project is to attempt to gain further understanding of how skilled and less skilled readers differ in relation to the most foundational elements of the task. A comprehensive study of effective preview is particularly appropriate for two reasons. First of all, as discussed earlier, the current empirical base is ambiguous about the importance of preview size to skilled sight-reading. It would seem desirable to investigate this measure in the context of a variety of notational complexities in order to provide a broad context for attempting to resolve these ambiguities. Secondly, if a regime of structural manipulations of performance material encompassing both skilled and less skilled readers is employed, further understanding may be achievable into the relative importance of sensitivity to musical patterning and basic perceptuo-motor factors in determining skill variation at the task.

It would also be beneficial to study visual perception sub-skills and motor sub-skills relevant to the complete sight-reading task. Motor ability has been especially underrepresented in the literature and therefore is a particular priority for attention. However, there is still a clear need for perceptual research, particularly into the issue of whether skilled and less skilled readers differ in their patterns of response, for example, the size of perceptual unit, or chunk, that they use for processing. Investigating the dependency of these sub-skills on musical patterning would also be a valuable undertaking considering that so little previous work has been carried out into this. Research into the sub-skills may also provide insight into their relative importance in determining skill at the complete sight-reading task.

The next chapter provides a broad introduction to the empirical work undertaken in this thesis to investigate the above issues.

Chapter 3

Introduction to research plans

3 Introduction to research plans

3.1 Introduction

In the light of the current state of knowledge about sight-reading as presented in the literature review, two related series of studies have been designed and undertaken:

1. Research into the complete sight-reading task, examining effective preview and sensitivity to musical structure in experienced pianists who were either skilled or less skilled sight-readers. Three studies have been carried out using a controlled preview methodology (described below), involving the performance at sight of monophonic, two-part and four-part materials.

2. Research into sight-reading related sub-skills using monophonic materials. Two studies were undertaken, one involving a visual perceptual processing task and the other an unrehearsed motor task, both employing the same set of skilled and less skilled subjects as in the first series of experiments. These studies were designed to gain insight into the origins of variation in the monophonic preview experiment data, but they are also of independent value.

Detailed methodologies for these two sets of investigations are found in Chapter 4 (controlled preview), Chapter 10 (perceptual sub-skill) and Chapter 11 (motor sub-skill). The principal aim of this chapter is to provide a general introduction to, and justification of, these experimental directions and strategies, organised under a single heading so that the individual parts can be clearly set within the overall research context. To avoid unnecessary duplication within the two methodology chapters, this chapter also contains discussion of two further issues relevant to both strands of the work: the use of tempo as a performance indicator and a description of the experimental participants.

3.2 Controlled preview research

3.2.1 Rationale and research goals

As has been discussed in the literature review, proponents of the patterning account view short-term memory storage as the principal constraint upon sight-reading skill. It is considered that skilled readers are able to significantly increase the efficiency of their short-term memory storage for notation together with the rate of encoding through their sensitivity to musical structure in the score, thus giving them access to greater preview with which to plan their output. Less skilled readers, on the other hand are seen as being less sensitive to musical structure, and thus unable to achieve equivalent memory, preview and thus performance levels. The discussion of prior research has raised questions about the empirical support for this account of skill in two areas. Firstly, there is no consistent evidence from the small number of studies that have been undertaken that skilled readers do make use of greater preview than less skilled readers. In the light of this, an alternative version of the patterning account may be appropriate, one in which skilled readers gain their performance superiority from the perception of structure within extents of preview equivalent to those used by less skilled readers. Secondly, to my knowledge, no work has been carried out clearly demonstrating that less skilled readers are any less sensitive to musical structure than skilled readers, or indeed that skilled reading is hindered by a lack of structural perception. The absence of any published research into how less skilled sight-readers perform with structurally disrupted notation means that the empirical base may equally point to skilled readers simply being significantly faster than less skilled readers at note processing generally, independently of structural influence. In this case, structure-related perception might only be of secondary importance in explaining skill difference, or perhaps not relevant to it at all, with the origins of skill to be found instead in basic perceptuo-motor factors and strategies.

There are therefore three basic research questions that need answering. Is the patterning account correct in proposing that skilled readers owe their ability to achieving larger preview than less skilled readers, mediated by their greater sensitivity to musical structure? Or might a modified patterning account be more appropriate in which the perception of patterning is still fundamental to skill but gained from a similar size of preview to that used by less skilled readers? Or should a primary role for patterning be rejected altogether in favour of an account based principally on a perceptuo-motor explanation of skill difference? Of course, it is quite possible that there is no simple answer to these questions and that a combination of patterning and perceptuo-motor explanations is needed for an overall explanation of skill. The questions, therefore, must not be considered as rigidly defining possible research outcomes, but rather as a helpful framework for the organisation of thinking in relation to experimental design and analysis. To attempt to answer the questions, methodology was developed for this thesis with the aim of gaining a greater foundational understanding of preview use amongst skilled and less skilled readers and also of the sensitivity of their sight-reading performance to musical structure.

3.2.2 Choice of methodology

A controlled preview methodology was considered appropriate for this part of the research for two main reasons. Firstly, it provides a means of quantifying effective preview without the need to have access to eye-tracking technology. Research using controlled preview has not, to my knowledge, been carried out within the domain of music sight-reading before, but it is commonly employed in the study of touch-typing. The study by Shaffer (1976) discussed in the literature review used it, and I will provide now a description of the approach in the context of his research. Shaffer's technology enabled him to limit the preview available to his skilled typist participants to a specific number of letters during their entire performance of a text sequence from a computer screen. Participants performed sequences under a variety of different preview size conditions ranging from a single letter to an entire line of text, and were required to type at the fastest speed at which accurate performance could be maintained. He found that the speed his participants were capable of increased with the number of letters of preview made available until an asymptote was reached, the preview size at that point marking the maximum level effective to the performance of that type of sequence (a measure that I will from now on call the 'maximum effective preview span'). His typists reached asymptotic performance levels of about ten characters per second with a preview size of eight characters. In addition, Shaffer manipulated the structure of the text stimulus in various ways. For example, he randomised word order and also randomised letter order within words. Within a musical implementation of the methodology, the influence of structural manipulations of musical sequences upon the maximum effective preview span is similarly open to exploration, something clearly important in relation to my research aims.

The advantage of this method of quantifying preview compared to Sloboda's eye-performance span approach is that it ensures a measure that is effective to performance. Sloboda (1985) mentions that his eye-performance span work was in fact an attempt to estimate the effective preview that a musical implementation of

Shaffer's methodology would have provided. The appropriate technology was not available to him, but he was clearly hopeful that before long, research using this approach would be carried out. 'No-one has yet published experiments with controlled preview in music, although the rapid development of computer music systems…makes such studies increasingly more feasible' (Sloboda, 1985, p. 71).

The second reason for choosing a controlled preview approach is because of the data that it provides at all the individual preview levels tested. These make possible a more complete understanding of the dependence of sight-reading ability upon preview, also enabling a detailed comparison of responses to different structural manipulations. Furthermore, the wealth of data for each subject across different conditions and preview sizes provides the potential for a comprehensive dissection and comparison of the performances of skilled and less skilled readers. The use of tempo as a measure of performance response, however, is a somewhat novel feature of this methodology in relation to music sight-reading research, and I will return to this matter later in the chapter.

Pianists were chosen as the participants for these studies because of the ease of effecting MIDI data transmission between electronic pianos and computer equipment/displays, something vital to the functioning of the controlled preview technology. Studying pianists also enabled the influence of notational complexity/task load upon performance variables to be explored through the use of both single and dual stave materials, providing a broader context for the analysis of skill difference. Some of the apparently contradictory evidence concerning skill and preview size from previous research relates to the performance of different complexities of material, and this aspect of the methodology may help to provide insight into this.

3.3 Perceptual sub-skill and motor sub-skill research

3.3.1 Rationale and choice of methodology

As discussed in the literature review, Sloboda considered there to be three ways to interpret the larger spans of apprehension achieved by the more experienced musicians in Bean's research (Bean, 1938). 'Does the superiority arise because experts have more rapid perceptual coding processes; or because they have better and more economical ways of storing what they perceive in memory; or because they have more efficient motor programs to organize the response? All the stages are plausible locations for the superiority.' (Sloboda, 1984, p. 224). To investigate the role of each of the sub-skills in the full task they need to be examined sufficiently in isolation from each other. To explore the area of perception and cognitive processing sub-skill, a pattern matching study similar in nature to the previously discussed work of Waters *et al.* (1998a) was undertaken. Such an experimental methodology is particularly appropriate to the study of perception-related behaviour because the motor component is minimal and very simple - pressing a key to indicate whether or not a match has been detected between the presented stimuli. A possible weakness with Waters *et al.'s* experiment was that there was no necessary requirement for musical encoding in the comparison procedure - participants were simply required to match sequences of notation. My own study represents a significant development upon their design, the comparison task involving a sequence of notation on a computer screen and a proposed transcription represented visually upon a graphic of a piano keyboard. Sequences of varying lengths were employed involving structural manipulations, in order to investigate how skilled and less skilled readers vary in the pattern and speed of perceptual processing, and how these are influenced by musical structure.

In relation to the study of motor sub-skill, the previous research of Lehmann and Ericsson (1996), discussed in the literature review, was limited in the conclusions it could draw about variation in sight-reading specific motor skill amongst able musicians because the methodology did not sufficiently isolate motor elements from the influence of standard notational input. In my motor sub-skill study an attempt at greater isolation has been made. Sequences of pitches and fingerings were presented using an animation displayed on a graphic of a piano keyboard, which participants had to memorise and subsequently reproduce in performance. With aims analogous to the perceptual sub-skill experiment, the sequences employed varied in length and involved structural manipulations.

3.4 Other issues

3.4.1 The use of tempo as a performance indicator

Within a controlled preview methodology, performance skill level is measured in terms speed of response. Whilst such a technique has been traditional within touch-typing research, and has also been used in the study of sight-reading sub-skills (Waters *et al.*, 1997), it appears to be quite novel in relation to the assessing of complete task sight-reading performance, where the typical method employed has been to quantify performance errors at a controlled tempo. There would seem to be no particular conceptual reason why sight-reading skill should not be assessed by tempo; the failure to use it as a performance indicator is possibly because musical activity, unlike touch-typing, is not normally associated with the idea of speed maximisation. However, as a gauge of raw transcriptional ability, tempo would seem to have much to commend it. Although it does not help to resolve measurement issues relating to the multidimensional nature of the sight-reading task (pitch finding, rhythm and expression) considered in the last chapter, it would seem to have particular benefits in relation to task demand issues, which will now be considered.

The literature review described how difficult it is to gauge precisely the skill of sight-readers across a range of ability levels from a single test piece performed at a set tempo. If the test piece is too easy, less skilled readers may well be distinguishable through the quantifying of performance errors, but the results for skilled readers may be subject to ceiling effects. If the test piece is too hard, the reverse may be the case, with the data of less skilled readers being difficult to distinguish due to floor effects. As mentioned earlier, Lee (2004) attempted to resolve this problem by employing a number of test pieces across a range of notational complexity/difficulty levels. Using a variety of task demands in this way was an important development in methodology enabling a closer match of test piece to individual participant ability than had previously been achieved. However, complexity of performance material is closely related to performance speed. For example, the need to process a greater number of parts clearly requires notes to be processed at a faster rate, if speed is kept constant. Taking this into account, the maximisation of tempo (supportive of accurate performance) required by the controlled preview methodology therefore enables an even closer matching of task to ability, with performance speeds attained in this way providing a quantification of skill at the task. This would seem a far more satisfactory way of measuring skill than quantifying errors, because its focus is upon what participants are actually able to accomplish, rather than the extent to which their performances break down under excessive task load. Clearly, participants will continue to make errors, though, and these remain important to assessing skill.

Given the novelty of a tempo-based methodology within a musical context, its validity and usefulness needs careful evaluation. For example, one objection to this approach is that one cannot be sure that participants are performing at maximum tempo levels. This issue will be considered further in Chapter 4. Also, parts of each stimulus sequence may vary in difficulty and elicit different performance speeds, something that would obviously complicate the study of rhythm reading. In view of this, the controlled preview research in this thesis has been confined to the study of isochronous materials.

3.4.2 Participants

As the literature review has shown, when studies fail to control adequately for technical proficiency, they are limited in the conclusions that they can draw specifically pertaining to sight-reading skill. It was decided in this research to focus upon skilled adult pianists, all either holding at least a teacher level diploma qualification from a music college or else a degree in music from a course which included a final year solo recital. It would have been preferable to control the performance level even more tightly, but this was not practicable in relation to finding sufficient participants able and willing to devote the considerable time and effort that this research demanded. It was felt important to use experienced pianists because, as will be seen in the methodology chapters, limitations in the technology employed, meant that more novice instrumentalists would have been more prone to producing corrupted data. These technological limitations also made it preferable to use adults, who would be more likely to undertake the tasks in a concentrated and committed manner.

Twenty-one participants took part in the research, responding to advertisements placed in the journal of the Incorporated Society of Musicians. Since they came from various parts of the country, this required travelling to the homes of most of the participants, although a small number carried out the studies in the Department of Music at the University of Sheffield. Whilst this may not have been ideal in terms of providing an identically controlled environment, it did mean that participants were at their ease and thus able to give valid accounts of their abilities. There were four male and seventeen female participants, varying in age from mid-20s to late 50s. Although it would have been preferable to have a greater control upon age, this was again not achievable in practice. None of the participants received payment.

Nineteen were involved in music professionally either as performers, accompanists or teachers, or indeed varying combinations amongst the three activities. Of the other two, one was a full-time homemaker and the other worked in a non-musical profession; both however were very active musically. Of the twenty-one in total, ten were self-categorised as skilled sight-readers, the other eleven as less skilled. The former categorisation represents readers who were confident enough in their reading skills for regular accompaniment to make up an important part of their musical lives. The latter categorisation represents readers who were unconfident in their sight-reading, evidenced by only rare or no involvement in accompaniment activity. These two skill groups form the basis for much of the statistical analyses carried out upon the research data. The decision to group participants by this method was taken in the light of the discussion in the literature review relating to the problems of sight-reading measurement. It was felt that current understanding about sight-reading was not sufficient to validly attribute a skill level to an individual on the basis of a standard 'test' piece of sight-reading. There was of course the risk that participants would wrongly categorise themselves, but there is no reason to assume that skilled musicians' assessments of their own reading ability level should not be valid, especially with the associated evidence of their level of practical engagement in sight-reading related activity. Furthermore, this research does actually require participants to carry out a large quantity of sight-reading with different structures and complexities of music, with analysis carried out on both an individual as well as a grouped basis. Therefore individual self-assessments can be checked against actual performances. It is hoped that the knowledge gained from this research will provide a greater understanding about how sight-readers of different abilities respond to different types of reading tasks, enabling more accurate and valid means of differentiating between skilled and less-skilled readers to be devised for future studies.

3.4.3 Structure of the remainder of the thesis

Chapters 4 to 9 focus upon the controlled preview research. Chapter 4 provides a more detailed description of the controlled preview experimental methodology begun in this chapter. After Chapter 5, which is a general introduction to the grouped statistical analyses of the preview research, Chapters 6, 7 and 8 involve specific analyses and discussions of the results of the monophonic, two-part and four-part notation studies respectively. Chapter 9 then undertakes an exploration of the data of individual participants from all three of these studies.

The perceptual and motor sub-skill studies are described and analysed in Chapters 10 and 11 respectively, with Chapter 12 examining individual participant data from these experiments in the context of the monophonic preview study results. Finally, Chapter 13 ties together the findings of both sections of the empirical work, and draws general conclusions from the research as a whole.

Chapter 4

Methodology of the controlled preview studies

4 Methodology of the controlled preview studies

4.1 Introduction

As mentioned in the last chapter, the controlled preview research in this thesis consists of three separate studies involving different, commonly used complexities of notation:

- Monophonic: treble clef stave and right hand only
- Two-part polyphonic: treble and bass clef staves, one note per stave/hand
- Four-part chordal: treble and bass clef staves, two notes per stave/hand

Because similar experimental procedures were employed in each case, it makes sense to describe these together in a single chapter. The results for the studies, however, are analysed separately in Chapters 6, 7 and 8 respectively. The aim of the work is twofold: first, to provide a foundational understanding of the dependency of sight-reading ability on preview across a range of task demands, and second, through the structural manipulation of experimental materials, to gain insight into the relative importance of sensitivity to musical structure and perceptuo-motor factors in explaining skill at the task.

With no previous research appearing to have measured effective preview in music sight-reading using Shaffer's (1976) controlled preview approach, discussed in the last chapter, it was necessary to design completely new technology for this set of studies. Starting with only rudimentary skills in computer programming, I developed appropriate software in the C language - a substantial undertaking running to 30,000 lines of source code. The program will be described in detail later in the chapter, but a brief introductory description would be useful. Essentially it is a musical equivalent of Shaffer's technology. A computer screen initially presents participants with the opening note or notes of a sequence, the number representing the preview size that performance is to be limited to, and which can be varied for different trials. This notation is displayed at the beginning of an otherwise blank stave, and as each note is played in order, beginning with the one on the furthest most left, a new note appears ahead of the one on the furthest most right. As each subsequent note is played, a new one continues to be added to the display until the end of the sequence is reached. Once displayed, all notes remain visible to participants.

As considered in the last chapter, a controlled preview methodology requires participants to perform at the maximum speed attainable for each preview size tested, this speed providing a measure of sight-reading skill for the particular experimental materials performed. For touch-typing, speed is the principal means by which performance level is defined, and typists are generally used to maximising their output. This is not the case with music sight-reading, though, and so to translate the controlled preview methodology successfully into a musical context requires participants to first to gain some practice at maximising their speed of response. Provided they are successful at this, the maximum effective preview span (the largest amount of preview necessary to sustain performance) can be quantified by requiring participants to perform sequences at a range of controlled preview sizes and identifying the one at which mean performance tempo reaches an asymptote. Data at smaller preview sizes offer the potential for understanding the manner in which insufficient preview constrains performance, and so a detailed picture of preview dependency can be built up for participants across the different materials performed.

4.2 Methodology

4.2.1 Introduction

No formal pilot study was undertaken. As the software was being developed it was regularly tested on a variety of musicians over an extensive period in the search for errors in coding and weaknesses in design. Such activity amounted to an informal piloting, and thus enabled judgements regarding the final methodology to be made: for example, the variety of preview sizes to be tested, the most suitable length for the sequences, the layout of the screen display and appropriate guidance for subjects in relation to producing experimentally valid performances.

4.2.2 Materials

4.2.2.1 General description of materials

To achieve insight into note-finding ability independent of skill at rhythm reading, each experiment - monophonic, two-part and four-part - focused exclusively on the sight-reading of isochronous pitch sequences. Crotchet-based notation was used for the monophonic and two-part studies, but the four-part materials were displayed using semibreves because programming variable tail lengths for crotchets would have been too time consuming. Although this was not ideal, it is not uncommon to find four-part compositions, for example arrangements of hymn tunes, presented in this way.

Individual sequences were 32 beats long (8 bars x 4 beats to bar), quite sufficient to create a satisfactory, albeit simple, musical architecture. When notated in an identical manner to Associated Board published sight-reading tests, sequences fitted onto a single line of music using a 17" computer monitor. Multiple lined sequences were not used in this study because of the interrupting effect of the eye having to move back to the left hand margin at the end of each line. The note range employed was limited to one leger line both above and below the staves, in order to constrain and standardise the purely visual complexity of the material to be read. The sequences did not employ a key signature; all sharps and flats were indicated by local accidentals. The reason for this was that during the piloting of the software it was found that when a key signature was used, participants sometimes forgot to play the required black notes in unstructured sequences because of the lack of a sense of tonal centre. Also, no set fingering was provided for the sequences. Fingering strategy may well play a role in sight-reading skill variation, and so free choice in fingering was given to allow this to be explored further from video documentation of performance (see below).

4.2.2.2 Factors studied

The sequences within each notational complexity (monophonic, two-part, four-part) were designed to facilitate the investigation of three factors that may lead to variation in performance: structure, key and preview size.

Structure

In order to explore the effect of musical structure on sight-reading performance, both structured and unstructured sequences were used.

Structured sequences

The structured materials displayed entirely conventional, triad-based tonal features, and were written in the style of simple folk and classical/romantic material. Melodies and bass parts were based on close triadic movement (to maximise local structural elements) with the standard range of embellishments used in a conventional manner i.e. passing notes, appoggiaturas and auxiliary notes. Closer part movement was typically required for the middle two parts of four-part material. The major constraint upon all this movement was that materials should lie comfortably under the hand.

The use of a 32-note sequence length divided into 8 bars of common time facilitated a balanced larger scale musical architecture with a variety of possible phrase structures. Characteristic melodic and harmonic progressions (the latter, of course, implied in the case of monophonic material) were provided to cadence points, and the forward movement was reinforced by a liberal use of melodic sequence, as well as some passing modulations (again implied in the case of monophonic material). Monophonic materials were limited to diatonic notes from the keys in which they were written. The harmonisation of modulations led to more use of chromatic content in the two-part and four-part sequences. Examples of these sequences will be provided following the discussion of key.

Unstructured sequences

Concerning the unstructured materials, by breaking the typical rules of tonal melodic and harmonic progression, sequences can be created almost entirely devoid of larger scale structure. However, when it comes to smaller scale patterning, the situation becomes more problematic. Almost any small group of notes is capable of some, albeit obscure, form of harmonic interpretation, and therefore it is clear that local level structure cannot be as effectively removed. All that can be done is to render such structure less obvious, by reducing triadic and passing note elements, and other more commonly met patterned content, for example, dominant seventh chords. This was achieved for the monophonic study with the aid of software that was designed capable of creating sequences using random note generation. Sequences were built up using an algorithm that rejected randomly generated notes that fell outside certain intervallic and directional constraints relating to those already included within the sequence. For example, no melodic intervals larger than an octave were allowed, and after an octave leap, the subsequent note would be required to return in the direction of the previous note; adjacent groups of either four ascending or descending notes were not permitted a range greater than an octave. Such constraints were imposed in order to create material that sat comfortably under the hand in a way similar to the structured sequences. Each sequence created was subsequently checked manually for ease of performance and any that were considered too awkward in comparison to their structured counterparts were rejected.

This software-driven design process was impractical for the more complex two and four-part unstructured sequence development. These materials were created manually by firstly reversing the previously composed patterned sequences, which removed any sense of their larger scale melodic and harmonic structure. Then, sufficient notes were altered by a small degree to disrupt more obvious local structure but not to significantly change the overall distribution of melodic and harmonic intervals. At the heart of this alteration process was also a concern to make the materials as comfortable to play as the structured sequences. Examples of these sequences are provided shortly.

<u>*Key*</u>

For all the three preview experiments, structured and unstructured sequences were provided at three levels of key, each based upon the notes of a diatonic major scale – C, G, and F major. Keys with more sharps or flats were not used in order to keep the tasks relatively simple. In the monophonic study, sequences were restricted entirely to notes diatonic to these three keys, but as was mentioned earlier, some chromatic notes were employed in the two-part and four-part studies. The rationale behind the use of the three keys, as opposed to a single key, was that this would enable the chosen range of the keyboard to be more fully explored. Structured sequences were always composed in the required key. Unstructured material was assigned a 'key' for monophonic materials because the random note generator only produced sequences of white notes. This was on the basis of whether adding an F#, a Bb or leaving it unchanged made it sit more comfortably under the hand. There was also a concern to make structured and unstructured sequences as equivalent as possible in terms of the numbers of black notes used. With two-part and four-part material, unstructured sequences remained in the 'key' of the structured sequence from which they were derived, and were similarly matched to this sequence in terms of black notes.
Key is not expected to have any significant influence upon performance, sequences in different keys being primarily viewed as replications within the structure condition. It was nevertheless considered important for key to be a separate condition in the statistical analysis so that its expected lack of influence could be verified.

Figure 4.1 and Figure 4.2 show structured and unstructured monophonic sequences in the 'key' of F. Figure 4.3 and Figure 4.4 show structured and unstructured two-part sequences in the 'key' of C. Finally, Figure 4.5 and Figure 4.6 show structured and unstructured four-part sequences in the 'key' of G. The complete set of sequences used for the monophonic, two-part and four-part studies are presented in Appendix 1. All sequences were independently evaluated (see below).

Figure 4.1
Example of structured monophonic sequence

Figure 4.2
Example of unstructured monophonic sequence

Figure 4.3
Example of structured two-part sequence

Figure 4.4
Example of unstructured two-part sequence

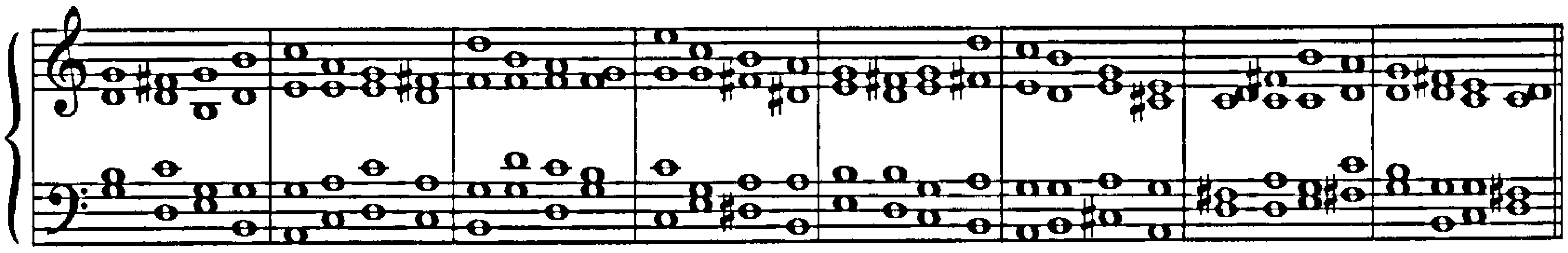


Figure 4.5
Example of structured four-part sequence

Figure 4.6
Example of unstructured four-part sequence

<u>*Preview size*</u>

The range of preview sizes used for each type of notational complexity was informed by informal piloting of the software. With evidence from this that maximum effective preview decreased in terms of beats from monophonic to two-part material and from two-part to four-part material, it was considered that employing the same range of preview levels for all types of notation would not have been an efficient use of experimental time. Therefore, controlled preview sizes from one up to seven beats were used for the monophonic study, one up to five beats for the two-part study, and one up to four beats for the four-part study.

As well as levels involving the limiting of preview, a further level for each study provided participants with unlimited preview of the entire sequence. This was employed to enable the validity of the controlled preview methodology to be tested i.e. if the asymptotic tempi achieved in controlled preview mode turn out to be significantly slower than tempi attained with unlimited preview, it might suggest that the technology (for example, the continual updating of the screen) was hindering performance. For each notational complexity, sequences were randomly allocated to a preview size so that for each preview size there were three structured and three unstructured sequences, one in each key. All participants performed all of the sequences organised according to this single random allocation. Whilst the robustness of statistical analysis would have been enhanced had there been a separate random allocation of sequence to preview size for each participant, this was not undertaken because it was considered that participants performing the same sequence/preview combinations would provide a rich resource for the study of skilled and less skilled reader fingering strategies.

4.2.2.3 Independent assessment of materials

All the materials allocated to a preview size had been previously evaluated by an independent adjudicator (a concert pianist, teacher and examiner) with regard to musical structure and how comfortably they lay under the hand. In relation to the former he was asked to categorise sequences as either:

- Very typical of tonally musical material
- Quite typical of tonally musical material
- Quite untypical of tonally musical material
- Very untypical of tonally musical material

With regard to comfort of performance, he was required to answer a single question in relation to each sequence: 'Appropriately fingered, does this sequence lie quite comfortably to very comfortably under the hand? Yes or No'. His evaluation of structure confirmed the design categorisation in all cases, with all structured materials judged very typical, and all unstructured materials very untypical, of tonally musical material. Also, all materials successfully met the ease of performance criteria – 48 monophonic, 36 two-part and 30 four-part sequences.

4.2.3 Apparatus

The experiments were performed on a Roland FP11 electronic piano, with weighted, touch sensitive keys that simulated a normal piano action. The presentation of the experimental materials was controlled from behind where subjects were seated, and thus out of their direct line of sight, using a Pentium 4 laptop computer. Sequences were displayed on a 17-inch flat screen monitor situated directly behind the piano on an adjustable height table, approximately 80cm in front of the participant's's face. This monitor was linked to the external VGA port of the laptop, and the MIDI OUT of the piano was connected to the MIDI IN of the laptop. Pitch and temporal performance data were recorded to the laptop's hard-drive. The notation was presented in black against a pale grey screen background, providing a comfortable, glare-free visual stimulus. A screenshot of the software is shown in Figure 4.7.

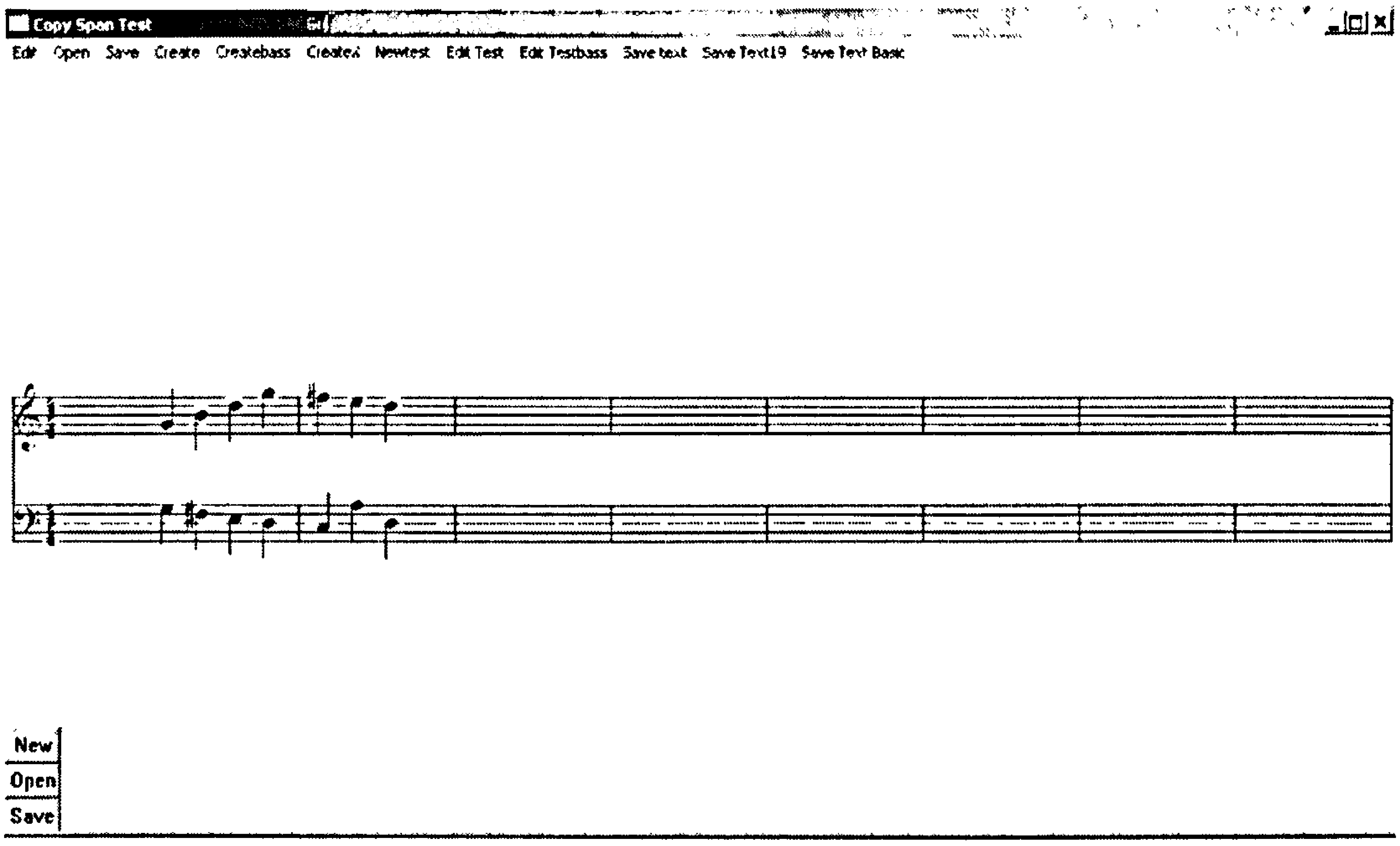


Figure 4.7
Screenshot of the controlled preview software showing part of a two-part sequence

During software development and piloting, some limitations of the planned technology became apparent. On the monophonic study, if participants played an extra note in order to correct an error, or hit two notes together by mistake, this resulted in an extra beat of preview being available for the remainder of the sequence, invalidating the subsequent data. The latter problem of accidental split notes was solved by adding code to the program that caused any note played within 100ms of the previous note to be ignored as far as screen updating was concerned. However, deliberate note-corrections were potentially more problematic, together with the associated issue of note omissions leading to a diminishing of preview size. The only way to deal with these was to ensure that subjects received training and sufficient practice prior to actual testing, and performed the sequences with a firm touch. This latter requirement means that there is no attempt to analyse dynamic data in the results of any of the studies. Split notes in the two-part sequences were dealt with using the same programming approach as for monophonic sequences. The issue of intentional corrections was rather less problematic with two-part material because two extra notes needed to be performed before an extra beat of preview was made available, one extra note having no effect. With four-part notation, split note programming was not a viable option because it depended upon there being no overlapping of parts, which was clearly not realistic to four-part sequence design. However, the facility was less necessary for this material, because four extra notes had to be performed before an extra beat of preview of preview was displayed.

In all three studies, therefore, participants had a certain margin for error concerning these issues, but there was still the potential for problems if performance was not carried out with sufficient care. During piloting it had been found that provided subjects were given sufficient training and time for practice, they invariably became comfortable with the required approach and performed reliably. Also, because the sequences were reasonably long, even if an occasional one became corrupted, in nearly all cases sufficient valid data was available for analysis. It is possible, though, that the required manner of performance may have acted to inhibit subjects from demonstrating their full tempo capabilities. This can be assessed by comparing the data for unlimited preview with those of the largest sizes of controlled preview, the former not involving any technological factors constraining performance.

A final technical limitation relates to the continual redrawing of the screen. During piloting it was found that a less-experienced musician who had to regularly look down at the keyboard during performance lost her place in the notation because of its continually changing appearance. Although it was felt that this would be largely resolved by focusing this research upon experienced musicians, a technological solution was also developed. An unobtrusive pale yellow mask was used to progressively cover the area of notation that had already been played, thus clearly indicating, if required, the next beat to be performed..

4.2.4 Procedure

The three controlled preview experiments were carried out in coordination with the two sub-skill experiments, the entire series normally carried out in one day, but occasionally requiring two half-days because of a participant's limited availability.

To facilitate the different schedules, the five experiments were divided into two sets, the first comprising the monophonic preview, two-part preview and the perceptual sub-skill studies, and the second, the four-part preview and the motor sub-skill studies. Within each set, a rest period of approximately 20 minutes was taken between studies, and when experimentation was carried out on a single day, a break of 1½ hours was taken between sets, typically for lunch. The first set of experiments always preceded the second set, and the within set order was always as defined above.

Each participant performed the sequences in a different random order. To mark the beginning of each trial a 3-2-1-start countdown was given, at which point the initial quantity of notes, representing the level at which preview would be controlled for that sequence, was presented on the screen. Participants were required to begin playing straightaway and to perform as quickly as possible to the end, without sacrificing accuracy of performance, and avoiding note omissions and repetitions. To help with this, they were informed that they could vary their tempo within a sequence depending upon the perceived difficulty of different sections. For monophonic and two-part material, legato performance was requested in order to ensure that a firm touch was achieved. With the four-part material, constraints upon fingering and the presence of repeated notes in adjacent chords meant that consistent legato performance was not possible. Participants were therefore instructed to use a detached, sustain-pedalled technique, a standard method for this type of music. For all studies, they were permitted to use any fingering that they felt to be appropriate, and asked to play with a firm touch, rather than focus upon musical expression. If the data for a sequence became corrupted close to the beginning of the performance, for example because of an extra notes having been played, performance was stopped immediately, and the entire sequence repeated at the end of the experiment. However, only data relevant to the previously unperformed section has been analysed. Such a procedure was essential so that the most efficient use was made of the available materials.

A short break of approximately 15 seconds was taken between trials whilst performance data was saved to the computer. For each study, the trials were undertaken in three groups, with participants allowed a rest of 5 minutes between groups. Each study took between 45 minutes to 1 hour to carry out, including practice time. The performances of all participants were recorded on video. Depending on the layout of a particular room, the camera was placed either two metres to the left or two metres to the right of the piano. At this distance, participants seemed to quickly forget about the presence of the camera, although there remains the possibility that the performances of some may have been negatively influenced by it. The tripod was set to a high position to enable a clear view of the fingers that were being used to play each key. The image took in the full range of keys required for the performance of the sequences, together with the head of the participant.

Prior to the formal experimentation of each study, participants performed practice sequences covering the different levels of the experimental conditions until they were confident with the performance requirements. As well as enabling more valid experimental performances, and ensuring less corrupted data, this also helped to minimise the influence of learning effects within the results. Participants were specifically encouraged to perform each type of practice sequence fast enough to induce errors, so that they gained a clear understanding of exactly what their performance limitations were. Despite this aspect of the practice trials, there is no independent guarantee that participants always performed at their maximum tempo, although it is possible that error rates may provide some indication of this.

Chapter 5

Introduction to results of controlled preview studies

5 Introduction to results of controlled preview studies

5.1 Introduction

This chapter provides first of all a summary of the independent and dependent variables employed in the controlled preview experiments, followed by an introduction to the nature of the grouped statistical analyses undertaken in the next three chapters (the monophonic analysis in Chapter 6, the two-part analysis in Chapter 7 and the four-part in Chapter 8). Finally, the pattern of controlled preview data that would provide support for each of the three accounts of sight-reading skill difference being tested by this group of studies is outlined, something that will help set the scene for later discussion.

5.2 Independent and dependent variables

For all three controlled preview experiments the independent variable conditions are the same. The only condition that changes in terms of the number of levels is preview size.

- Skill (2 levels: 10 skilled and 11 less skilled participants)
- Structure (2 levels: structured and unstructured)
- Key (3 levels: diatonic notes from C, G and F major)
- Preview size:

    Monophonic (8 levels: 1, 2, 3, 4, 5, 6, 7 beats and unlimited preview)

    Two-part (6 levels: 1, 2, 3, 4, 5 beats and unlimited preview)

    Four-part (5 levels: 1, 2, 3, 4 beats and unlimited preview)

In the monophonic study, each participant played 48 sequences, one sequence in each key for each structure level at the 8 preview sizes. On the two-part and four-part studies the number of sequences was 36 and 30 respectively, the difference resulting from the smaller number of preview size levels.

The principal dependent variable quantified by the studies is inter onset interval - the duration between successive key presses. It is not appropriate to use individual inter onset interval data points for the statistical analyses because it cannot be guaranteed that they are sufficiently independent of each other. This can be understood by considering a scenario where there is little variation in inter onset interval within the performance of a sequence. Whilst such a lack of variation may well point to a participant requiring an equivalent amount of time to respond to each of the notes, it may equally represent performance undertaken at a set tempo, with not all notes necessarily being responded to at a peak level. With the individual data points not usable, the analysis has been carried out in terms of the mean tempo of each sequence performance rather than the mean inter onset interval. This was principally in order to make the data and graphs more intuitive for the reader. For example, it would seem more intuitive to conceive of tempo increasing with greater levels of preview, rather than inter onset interval decreasing. The other dependent variable quantified was note errors - performed notes that do not match sequence notation. The error rate for the studies was very low, and so no statistical analysis has been carried out on these particular data.

Participants' initial perception of the opening notation for each sequence involved no accompanying motor activity, something that may have made this material easier to process and perform. Also, the playing of the final notes of each sequence caused no further notation to be added to the stave, meaning that motor activity here was accompanied by less perceptual demand, again possibly leading to easier performance. To remove these potential effects from the data, the first 7 beats and last 6 beats of the monophonic sequences, the first 5 beats and last 4 beats of the two-part sequences, and the first 4 beats and last 3 beats of four-part sequences have not been included in the calculation of the mean tempo. These figures represent the beats potentially subject to the effect at the maximum level of preview size for each complexity of notation. This degree of pruning might not have been necessary for smaller preview sizes, but it was considered important to treat the entire data for each analysis in a consistent manner. The mean tempi used for the analyses are therefore based on 19 inter onset intervals for monophonic sequences, 23 for two-part sequences and 25 for four-part sequences. Occasional performances produced fewer data points as a result of corruption (discussed in the last chapter), but provided that there were at least 10 adjacent inter onset intervals obtained, the results were deemed permissible. The number of completely corrupted and hence unusable sequence performances was very low: 14 out of the 1008 monophonic trials, 5 out of the 756 two-part trials and 6 out of the 630 four-part trials.

5.3 Description of statistical analysis

The same basic statistical analysis has been carried out on the grouped data of each of the three controlled preview studies: a 3-factor within subjects (preview size, structure, key), 1-factor between subjects (skill) repeated measures ANOVA. Following this, in Chapter 9, there is an examination of individual subject data from the three experiments. For valid ANOVA, the data are required to comply with certain theoretical assumptions underlying the statistical model. The individual levels of the 'between' factor, in this case skill, should demonstrate normality and homogeneity of variance. The assumption of homogeneity of variance is adequately met by the distributions in all the experiments. However, as is typically the case in studies measuring reaction-type responses, the distributions are positively skewed and hence not ideal in terms of ANOVA's normality requirements. However, ANOVA is typically robust in the face of such non-normality, though, and considering that the distributions of the skill groups' data in each experiment are similarly shaped, they can be deemed acceptable.

A further assumption specifically of repeated measures ANOVAs is that the data display sphericity, that is, the variances of the differences between the levels of the repeated measures factor are required to be equal. This is typically not the case in the three studies and has therefore been compensated for by applying the Greenhouse-Geisser epsilon adjustment, the most conservative approach, to the degrees of freedom of the relevant effects. Wherever this epsilon adjustment has been used, the statistical result is presented in a slightly different format to usual: for example, $F_{7, 133}$ [1, 19] = 60.62, $p < 0.0001$. The numbers in subscript here represent the unadjusted degrees of freedom, and those in square brackets the degrees of freedom after multiplication by the epsilon adjustment, which in this case has the value 0.143 (to 3 decimal places). A standard significance level of 5% has been employed throughout the analysis.

Simple effects and simple interactions have been calculated using smaller scale within-subjects and between-subjects ANOVAs, following the advice of Howell (2002) in relation to repeated measures designs. Planned comparisons have been undertaken in each study using linear contrast analyses (t-tests) to identify the preview size at which asymptotic performance level is reached for each skill group. Only a very small number of post hoc comparisons have been undertaken: none on the monophonic study and two on both the two-part and four-part studies. The significance level has been held at 5% for this small number of investigations because it is considered that the robustness of the mean tempo data, together with the number of replications (the three levels of key) is a sufficient guard against Type 1 errors (accepting an alternative hypothesis when the result is attributable to chance).

5.4 Accounts of skill to be tested

To provide a context for the statistical analysis, I will outline the pattern of results that would be expected for each of the three accounts of sight-reading skill difference that the controlled preview studies have been designed to test. As mentioned in Chapter 3, these accounts are not to be considered rigid definitions with necessarily an independent reality; it is possible that an overall understanding of the task may require a combination of some of their elements. They are primarily to be viewed as useful constructs to facilitate the logical organisation of the analyses.

The patterning account

For each study, the skilled group should perform significantly faster than less skilled group by making use of greater preview, but principally only in relation to structured material, indicating their more efficient levels of memory storage to be structure dependent. The less skilled group should perform similarly on both structured and unstructured sequences, although some limited sensitivity to structure may be expected. For both types of material, the performance of the two skill groups should be fairly similar for levels of effective preview that they have in common, indicating their generally equivalent basic perceptuo-motor abilities. There may well be some sensitivity to structure displayed within this range, but the performance gains achieved by the skilled participants with structured notation should be achieved primarily as a result of their use of greater preview.

The modified patterning account

This account is a modification of the patterning account to attempt to explain research evidence showing skilled readers to have similar effective levels of preview to less skilled readers (Gilman & Underwood, 2003). As with the patterning account, skill should primarily be mediated through the perception of musical structure, but with the mechanism involving skilled participants processing structured material more quickly than less skilled participants without using significantly larger amounts of preview.

The perceptuo-motor account

The superior performance of the skilled group would be expected to lie primarily in their being faster at transcribing notes generally than the less skilled group, and therefore the former should perform more quickly than the latter at all levels of preview, with both structured and unstructured materials. It is possible that skilled participants may make additional performance gains by being able to use larger amounts of preview. There may also be some limited variation between the performance of structured and unstructured materials in terms of both preview use and tempo, with the small amount of previous research suggesting that the skill groups will differ little from each other in this regard.

For all three accounts, it would be expected that the dependency on structure would become more pronounced as the number of parts, and hence the task demand, increases.

Chapter 6

Monophonic controlled preview study

6 Monophonic controlled preview study

6.1 Introduction

6.1.1 Résumé of methodology

The 10 skilled and 11 less skilled participants used their right hand to sight-read monophonic, treble stave sequences, each consisting of 32 crotchets divided up into 8 bars of common time. The following regime of independent variable conditions was employed:

- Structure - 2 levels: structured and unstructured
- Key - 3 levels: diatonic notes from C major, G major and F major
- Preview size - 8 levels: 1, 2, 3, 4, 5, 6, 7 beats, and unlimited preview

For each of the 8 levels of preview size, there were structured and unstructured sequences in all 3 keys making 48 sequences in all. Participants were required to sight-read all sequences as fast as they could comfortably manage without sacrificing care and accuracy in their performance, and pitch and timing data were recorded for performances. A mean tempo in beats per minute has been calculated for each sequence providing a total of 1008 data points (48 trials x 21 subjects) for the statistical analysis, which comprises a 3-factor within subjects (structure, key and preview size), 1-factor between subjects (skill) repeated measures ANOVA.

6.1.2 Aims of the analysis

Although the patterning account is the dominant explanation of sight-reading skill difference within the domain, evidence from previous research is also consistent with other accounts. The patterning account considers the principal source of skill difference to be skilled readers' greater sensitivity to musical structure within the score, enabling them to process greater quantities of preview at a faster rate than less skilled readers. However, there is evidence that less skilled readers are able to use similar quantities of preview to skilled readers, indicating either a modified patterning account, in which skill is still mediated primarily by structural perception but independent of preview size, or a perceptuo-motor account, in which skill is principally the result of perceptuo-motor factors. The main aim of the analysis is therefore to test the validity of the patterning account, and to consider the extent to which the other two accounts are supported by the data should the patterning account be found wanting. A more detailed summary of the results that would provide support for each of the three different accounts is found at the end of Chapter 5.

6.2 Results

6.2.1 Main effects

To provide a general context for the analysis, the main effects will be considered to begin with. Only two out of the four main effects, skill and preview size, are at a statistically significant level. The effect of skill is highly significant ($F(1, 19) = 28.3$, $p = 0.0004$), the skilled group achieving an overall mean tempo of 244 beats per minute (SD = 76) and the less skilled 182 beats per minute (SD = 48). Preview size is similarly highly significant ($F_{7, 133}[1, 19] = 238.27$, $p < 0.0001$) and the data relevant to this effect is presented in Table 6.1 and Figure 6.1.

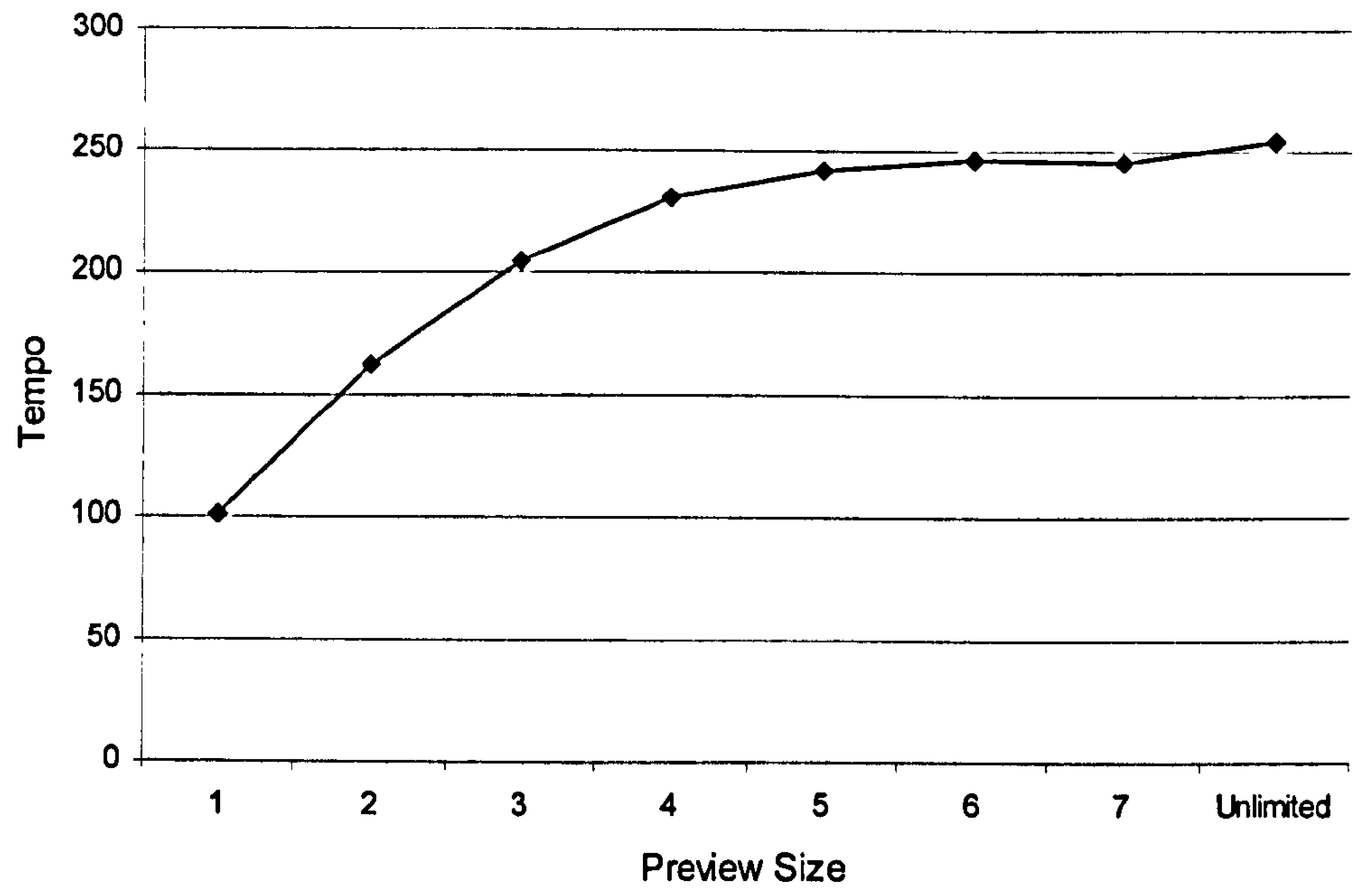


Figure 6.1

Mean tempo (beats per minute) at levels of preview size (beats) (combined across skill, structure and key)

Table 6.1

Mean tempo (beats per minute) at levels of preview size (beats) (combined across skill, structure and key, SD in parentheses)

Preview Size	1	2	3	4	5	6	7	Unlimited
Tempo	101 (13)	163 (22)	205 (34)	231 (49)	242 (52)	247 (61)	246 (60)	254 (70)

For combined skill groups, the distribution is quadratic in shape, mean tempo increasing steeply with preview to begin with, then gradually levelling out to an apparently stable asymptote. This distribution is comparably shaped to those obtained in touch-typing controlled preview research (Salthouse, 1986), demonstrating the similar dependence of sight-reading and typing upon overlapping processing operations to achieve rapid output. The main effect of structure is non-significant ($F (1, 19) = 1.07, p = 0.31$), however it does not necessarily follow from this that the performances of the skill groups were not subject to structural influences. This issue will be discussed later in the chapter. The main effect of key was also non-significant ($F_{2, 38} [1, 19] = 0.04, p = 0.84$), a result indicating that the use of different keys has, as planned, provided the desired variety of content for sequences without becoming a source of unwanted variation

6.2.2 Interactions and simple effects

This part of the analysis is divided into three sections. Firstly, there is a brief introduction to the pattern of preview use of the individual skill groups. Secondly, consideration is given to the manner in which structure is influential within the data. Finally, the issue of preview is revisited in more detail, taking into account the findings of the structural analysis. In particular, the maximum effective level of preview is determined for the groups.

<u>6.2.2.1 Pattern of preview use of the skill groups</u>

The skill x preview interaction is highly significant ($F_{7, 133} [1, 19] = 23.9, p = 0.0001$), and the separate preview distributions of the two skill groups, both similar in shape to the combined distribution in Figure 6.1, are displayed in Table 6.2 and Figure 6.2. The interaction is clearly visible in Figure 6.2. The less skilled group perform only 9 beats per minute slower than the skilled group at 1 beat of preview (skilled = 105 beats per minute, less skilled = 96 beats per minute), but then make increasingly less use of each extra beat than the skilled readers. From a merely visual inspection, the skilled readers appear to employ a larger quantity of preview in achieving asymptotic performance, and their maximum tempo of 309 beats per minute, recorded at unlimited preview, is 101 beats per minute (49%) faster than the less skilled maximum of 208 beats per minute, recorded at 5 beats of preview. Maximum tempi are approximately three times faster than 1-beat preview performance for the skilled group, and twice as fast for the less skilled group. There are two particular statistical investigations required here: firstly, to determine at what preview size the between-group tempo difference becomes significant, and secondly, to quantify the maximum effective preview span for each skill level i.e. the preview size after which no further significant gains in performance are made. Before undertaking these, though, consideration needs to be given to possible effects of structure upon the two distributions.

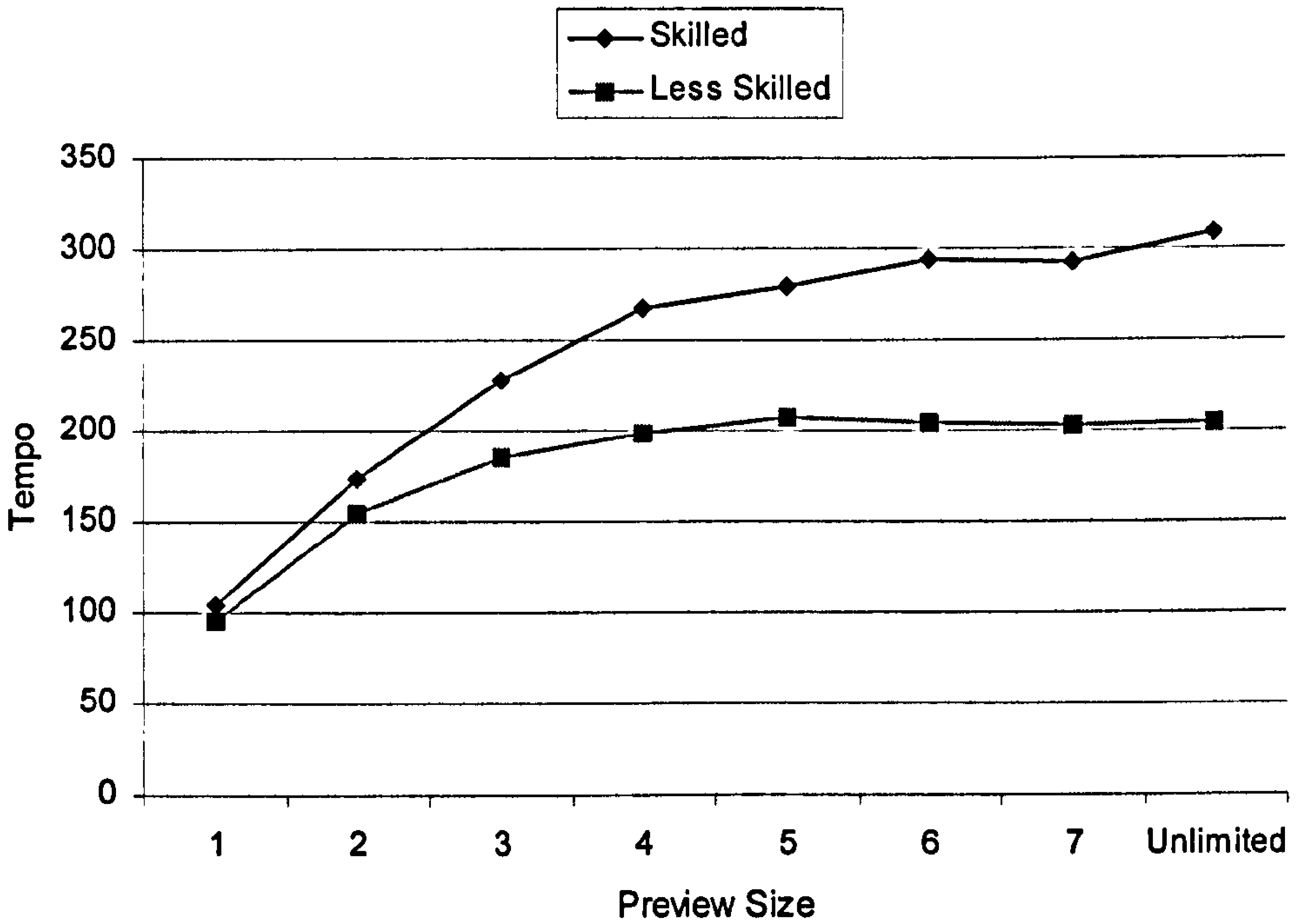


Figure 6.2
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key)

Table 6.2
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key, SD in parentheses)

Preview size	1	2	3	4	5	6	7	Unlimited
Skilled	105 (14)	174 (21)	228 (25)	267 (32)	280 (35)	294 (42)	293 (44)	309 (56)
Less skilled	96 (11)	154 (18)	185 (27)	198 (36)	208 (40)	204 (39)	203 (35)	205 (36)
Skilled minus less skilled	9	20	43	69	72	90	90	104

6.2.2.2 Structure and its interactions with skill and preview

There is a weakly significant skill x structure interaction (F (1, 19) = 4.21, p = 0.05), and the data relating to this show that the skilled group is on average slightly faster at performing unstructured than structured material, the reverse being the case for the less skilled group. For the skilled readers, at combined conditions of preview and key, the mean tempo of the structured sequences is 243 beats per minute (SD = 74) and of the unstructured sequences 245 beats per minute (SD = 77). For the less skilled readers the respective figures are 184 beats per minute (SD = 46) and 180 beats per minute (SD = 51). The simple effect of structure is not significant for the skilled group (F (1, 9) = 0.65, p = 0.44) but approaches significance for the less skilled group (F (1, 10) = 4.12, p = 0.07). The preview x structure and the preview x structure x skill interactions are both non-significant ($F_{7,133}$ [0.47, 8.87] = 2.56, p = 0.14 and $F_{7,133}$ [0.47, 8.87] = 0.53, p = 0.35 respectively) indicating that the influence of structure is consistent across the preview range for both skill groups

6.2.2.3 Analysis of the skill groups' pattern of preview use

In view of the non-significance of the main effect of structure and of the interactions just discussed, there is no justification for any statistical investigation of the separate levels of structure with regard to preview use. The two statistical investigations proposed earlier have therefore been undertaken using combined structure levels.

Firstly, using 1-factor between-subjects ANOVAs, the simple effect of skill at 1 beat of preview is found to be non-significant, but at 2 beats of preview a significant result is obtained ($F(1, 19) = 7.58$, $p = 0.01$). Therefore, the significant main effect of skill is found to be relevant only to preview levels of 2 beats and above. Secondly, planned linear contrasts have been used to quantify maximum effective preview spans for the two skill groups. The preview size at which asymptotic tempo is reached could have been quantified by comparing pairs of tempo values at adjacent preview sizes, for example, 2 and 3 beats, 3 and 4 beats, 4 and 5 beats and so on, until a statistically non-significant result was obtained. However, a weakness of this method is that an idiosyncratic tempo value at a single preview size may mislead the analysis. For example, the tempo for less skilled subjects at 5 beats of preview is faster than at higher levels (see Table 6.2). Simply comparing the values for 4 beats and 5 beats, therefore, may possibly lead to an overstating of the importance of the 5^{th} beat of preview. Comparing the tempo value of a single preview size with the mean of all the larger preview sizes, on the other hand, helps to spread the effect of eccentric data, increasing the likelihood of a valid analysis. The appropriate weightings for the linear contrast analyses together with the results are presented in Tables 6.3 and 6.4. The findings are that the skilled readers make no statistically significant performance gains beyond 6 beats of preview, the equivalent figure for the less skilled readers being 4 beats of preview.

Finally, Table 6.5 presents again the data of Table 6.2, but combines the results at maximum effective preview sizes and above. As well as demonstrating the comfortable maximum tempi that can be achieved at each level of preview, these data can also be viewed as providing an indication of the level of preview that participants actually need to attain particular tempi. For example, Table 6.5 shows that the less skilled readers can perform up to 185 beats per minute with 3 beats of preview. From this it can be inferred that to perform significantly beyond such a tempo they will need 4 beats of preview. This manner of reading the data from this table will prove useful at certain points in later discussion.

Table 6.3

Linear contrast analyses to determine maximum effective preview span: skilled group

Preview size	2	3	4	5	6	7	Unlimited	
Weighting a	-6	1	1	1	1	1	1	$t(63) = 16.28, p < 0.0001$
Weighting b	0	-5	1	1	1	1	1	$t(63) = 9.43, p < 0.0001$
Weighting c	0	0	-4	1	1	1	1	$t(63) = 4.04, p = 0.0001$
Weighting d	0	0	0	-3	1	1	1	$t(63) = 2.67, p = 0.001$
Weighting e	0	0	0	0	2	1	1	$t(63) = 0.90, p = 0.37$

Table 6.4

Linear contrast analyses to determine maximum effective preview span: less skilled group

Preview size	2	3	4	5	6	7	Unlimited	
Weighting a	-6	1	1	1	1	1	1	$t(70) = 10.66, p < 0.0001$
Weighting b	0	-5	1	1	1	1	1	$t(70) = 4.31, p < 0.0001$
Weighting c	0	0	-4	1	1	1	1	$t(70) = 1.42, p = 0.16$

Table 6.5

Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key), with results at and above maximum effective levels of preview combined

Preview size	1	2	3	4 (skilled) 4-plus (less skilled)	5	6-plus
Skilled	105	174	228	267	280	299
Less skilled	96	154	185	204	----	------
Skilled minus Less skilled	9	20	43	63		

6.2.3 Performance Errors

The performance error data are combined figures relating to two kinds of error: notes misplayed and notes omitted. The number of errors for both groups is very low. On average, each skilled participant performed 4 errors throughout the entire 48 sequences (0.45% error rate), and each less skilled participant performed 14.7 errors (1.6% error rate).

6.2.4 Summary of Results

The skilled group performs significantly faster than the less skilled group at all levels of preview except for 1 beat. For each skill group, the effect of preview is highly significant, producing distributions that are quadratic in shape - tempo initially increasing sharply with preview size, followed by a gradual tailing off towards an asymptote. There is a significant skill x preview interaction, with the difference in tempo between the skill groups only small (9 beats per minute) at 1 beat of preview, but rising to 63 beats per minute by 4 beats where the asymptotic speed of the less skilled group is reached (204 beats per minute). Skilled readers continue to make further performance gains with greater levels of preview, reaching their asymptote (299 beats per minute) at 6 beats. The pattern of behaviour displayed is independent of any significant structural effect, except for a skill x structure interaction resulting from skilled subjects performing slightly faster with unstructured material, and less skilled subjects performing slightly faster with structured material. This interaction does not lead to a significant simple effect of structure for either skill group.

6.3 Discussion

6.3.1 Introduction

I will begin with an introductory comparison of the preview use of the skill groups, followed by two sections evaluating whether the evidence points to skill difference at the task being primarily due to perceptuo-motor factors or to sensitivity to patterning. In the final section the issue of factors limiting performance is discussed.

6.3.2 Introductory comparison of the skill groups' preview data

At the heart of the patterning account is Sloboda's contention that his skilled readers depended upon their larger monophonic eye-performance spans (skilled = 7 notes, less skilled = 4 notes) in order to perform more accurately and fluently at the same tempo as his less skilled readers (Sloboda, 1974). Although my participants' maximum effective preview spans (skilled = 6 notes, less-skilled = 4 notes) are very similar in size to Sloboda's measures, it needs to be borne in mind that they were recorded with my skilled group sight-reading considerably faster than my less skilled group. In fact, my skilled group did not need greater preview to outperform their less skilled counterparts: their performance at 3 beats of preview (228 beats per minute) is faster than the asymptotic speed of the less skilled group, achieved at 4 beats of preview (204 beats per minute). At 4 beats of preview, the skilled readers themselves achieved 267 beats per minute, representing a tempo advantage over the less skilled readers of 31%. The skilled asymptotic tempo of 299 beats per minute, reached at 6 beats of preview, represents only a further 12% increase upon this 4-beat level. Therefore, there is no evidence that the roots of the skill difference displayed in this study lie in skilled readers perceiving greater quantities of preview than less skilled readers, as proposed by the patterning account. Larger preview simply enables skilled participants to enhance to a small extent what is already a considerably superior performance within the less skilled subjects' preview range. Clearly, being based upon the data from only one study this finding needs to be viewed cautiously. It also may not be representative of monophonic reading in general considering the lack of rhythmic variation in the materials. That having been said, though, it is certainly consistent with other evidence presented in the literature review questioning of the importance that has been attached to preview size in explaining skilled monophonic reading. For example, it will be remembered that one of Sloboda's skilled eye-performance span subjects found the sight-reading of disrupted material no harder than structured material, despite performing the former with a significantly smaller eye-performance span (Sloboda, 1985).

Although the source of skilled sight-reading in this study lies, contrary to the expectations of the patterning account, in the skilled group's more efficient transcribing of preview levels shared with the less skilled group, this does not necessarily deny a primary role for the perception of structure in determining skill. Clearly, however, the mechanism by which structure drives performance needs to be reconsidered to make such a role consistent with the preview data. Specifically, it is necessary for skilled subjects to have derived most of their structure requirements from a significantly narrower notational context than that proposed by the patterning account itself. In other words, if the perception of structure does play a primary role in sight-reading skill difference, it is the modified version of patterning account that provides the more appropriate explanation.

Before evaluating the modified patterning account and the perceptuo-motor account in relation to the data, some consideration needs to be given to why Sloboda's eye-performance span measures show his skilled readers using considerably greater preview than his less skilled readers whilst performing at a similar speed to them, whereas my skilled group outperformed my less skilled group's maximum tempo in the context of less preview. One possibility that has already been discussed in the literature review is that Sloboda's skilled eye-performance span figure may not be entirely effective in nature i.e. it may include a level of perception that is surplus to performance requirements. However, another possible explanation is that his measure and my own are indeed both effective, but quantify the preview requirements of different skill components. Although my skilled group only needed 3 notes of preview to outperform the less skilled group, such a level of preview is clearly insufficient for the planning of fluent fingering and musical expression. With performance at the same tempo under ecological conditions, preview no doubt would normally be extended beyond this rather artificial level in order to effect a more appropriate degree of motor planning and control. Therefore, my 3-note preview figure may be effective in 'raw' note-finding terms, but Sloboda's 7-note eye-performance span may more appropriately express the extra preview required by skilled readers to produce musically fluent performances. As was made clear in the literature review, Sloboda developed the patterning account assuming that the full eye-performance span of his skilled readers was essential to note identification, and therefore the above discussion suggests that his account may have been based upon a misinterpretation of the meaning of his measure.

6.3.3 Support for the perceptuo-motor account within the data

One possible reason for the absence of any significant structural effect within the data is that the monophonic sequences were not demanding enough for either skill group to need structural perception to help them achieve maximal performance. In other words, the perception of notes without reference to musical meaning was quite sufficient to provide all the necessary processing speed and memory storage, with skilled readers owing their performance advantage to their superior perceptuo-motor functioning. Such an account makes no assumptions about how aware the skill groups were of structural content; the ability to perceive structure is simply not considered relevant to output. Of course, it is also possible that the material was sufficiently demanding for structure to be of use in performance, but that there was not enough difference between structured and unstructured sequences for an effect to be elicited. This would seem unlikely to have been the case, though - it is difficult to see how any more structure could have been added to the structured sequences, or how more could have been removed from the randomly generated unstructured sequences.

The data would appear to be entirely consistent with the first explanation. The lack of a significant difference in performance between the groups at 1 beat of preview indicates that the skilled and less skilled readers did not vary in terms of basic reaction level responses to notation. This is in line with the findings of Lee (2004), that there is no significant correlation between reaction time and sight-reading ability.
At 2 notes preview and beyond, the increasing tempo of the skilled group compared to the less skilled group can be readily accounted for in perceptuo-motor terms, with skilled readers being more efficient at processing multiple notes and/or performing overlapping perceptual and motor processing operations. Clearly, no conclusions about specific causes of the performance variation between the groups can be drawn from these data alone. Less skilled readers may have been using the same strategies and memory representations as skilled readers but at a slower rate. On the other hand, their strategies and representations may have been different and less efficient, or perhaps they had difficulty with a particular area of sub-skill. The fact that the less skilled group performed more errors than the skilled group suggests the latter explanations to be more likely. The sub-skill studies in the second part of this thesis may shed further light upon this issue.

6.3.4 Support for the modified patterning account within the data

There are at least two other possible reasons for the absence of a structural effect. Firstly, sufficient structure may have remained within the unstructured sequences to enable a highly structure-sensitive skilled group to maximize their performance. Because the structured sequences can be considered to have subsumed larger scale patterning, the lack of an effect of structure indicates that predictive or inferential mechanisms did not play a significant role in performance. Secondly, the less skilled group may have been highly insensitive to structure, confined to processing all notes on an individual basis, and achieving tempo gains from greater preview entirely through parallel processing efficiencies. Taken together, these interpretations offer the potential for a primarily patterning driven account of skill. According to this explanation, the equivalent results of the skill groups at 1 beat of preview, where structure can obviously have no effect, is representative of their similar perceptuo-motor ability, with gains made by the skilled readers at increasing preview levels attributable to the growing notational context providing a greater opportunity for structural perception and structure driven output efficiencies.

There are questionable aspects to such an account of skill though. First of all, whilst it is conceivable that with a large enough context of preview structure-sensitive readers might glean sufficiently meaningful patterning from unstructured sequences to maximize their performance, it would seem doubtful that the small preview sizes shared with the less skilled group would be capable of providing this. For example, any structural perception obtained from even structured material would be necessarily ambiguous at 2 beats of preview and therefore seem unable to account for the 20 beats performance advantage of the skilled group here (skilled = 174 beats per minute, less skilled = 154 beats per minute (see Table 6.5)). Also, the unstructured sequences were largely devoid of any adjacent triadic or passing note content, making it difficult for sensitivity to patterning to account for the even greater advantage of the skilled group at 3 beats of preview (skilled = 228 beats per minute, less skilled = 185 beats per minute). Another problem with this patterning based explanation of the data is the required assumption that the less skilled participants are highly insensitive to musical structure. Considering that this group are otherwise able musicians it would seem improbable that this is the case.

Overall, therefore, the evidence would suggest the perceptuo-motor account of the data to be the more plausible of the two. The case in support of a primary role for structure in explaining skill is considerably weakened by the narrower notational context for structural perception demanded by the modified version of the patterning account. However, clearly no firm conclusions can be drawn on the basis of this study alone.

6.3.5 A discussion of factors limiting performance

Sloboda's skilled readers experienced a significant decrease in their eye-performance spans when they sight-read structurally disrupted notation (Sloboda, 1977). He considered this as evidence that monophonic reading ability was limited by memory storage availability i.e. that skilled levels of reading required the extending of this storage, and hence preview, through chunking mechanisms. His results contrast, however, with the experience of my skilled participants who achieved equivalent maximum effective preview spans on both structured and unstructured sequences. These results suggest that the memory demands of monophonic sight-reading are not sufficient to require the employment of chunking mechanisms, and therefore that some other factor or factors within the perceptuo-motor system were responsible for preventing both skill groups from achieving further performance gains. One possible explanation for the difference between my results and those of Sloboda is that his eye-performance span measure overstated the explicit memorisation of notation needed for skilled levels of single-line reading. It could also be that my participants were able to continually refresh their memories from the score, whereas the recording of Sloboda's eye-performance span required unprompted memory that might have been more prone to decay with unstructured material.

It was discussed in the literature review how time-constrained measures of perceptual uptake may possibly understate short-term memory storage. If insufficient time is made available to experimental participants, for example, by the performance tempo being set too fast, they may only be able to partially fill their short-term memory buffer, meaning that measures obtained for skilled and less skilled readers may simply represent their relative rates of encoding. The data for the current study are consistent with such an idea: my skilled group's asymptotic processing of notes was approximately 50% faster than that of my less skilled group (skilled = 299 beats per minute, less skilled = 204 beats per minute) and they achieved this in the context of 50% larger preview use (skilled = 6 beats, less skilled = 4 beats).

Clearly, no firm conclusions can be drawn about factors limiting performance from this discussion. However, the dual stave study of Gilman and Underwood (2003) discussed in the literature review has indicated that skilled and less skilled readers may be capable of storing similar quantities of notation in short-term memory, levels that are larger than my skilled reader monophonic maximum effective preview spans. Therefore, further light may be shed upon these issues from the two-part study data in the next chapter, and so this particular discussion will continue there.

6.3.6 Summary

The specific mechanism proposed by the patterning account to explain skill variation between skilled and less skilled readers is not supported by this study's data. The root of skill difference on the monophonic preview task was found to lie not in the skilled group's use of greater preview than the less skilled group, but rather in their more efficient use of preview amounts within the less skilled readers' effective range of 4 beats. Skilled readers were only able to use their extra 2 beats of preview to achieve small further increases in performance tempi. A modified patterning account was proposed to explain, from a structure based perspective, the greater processing efficiency of skilled readers with smaller levels of preview. However, the perceptuo-motor account would seem to provide a more plausible explanation of the data than this. In particular, it is difficult to see how unstructured sequences could have delivered sufficient unambiguous structure at low preview sizes to account for the significantly faster skilled group performances. Finally, there is no evidence that either group's performance was limited by short-term memory capacity. Other, as yet unspecified, perceptuo-motor factors may therefore have been acting to prevent further performance gains.

Chapter 7

Two-part controlled preview study

7 Two-part controlled preview study

7.1 Introduction

7.1.1 Résumé of methodology

The 10 skilled and 11 less-skilled participants performed at sight, with hands together, 36 two-part sequences, each sequence comprising 2 parallel staves of 32 crotchets (one part treble clef, the other bass clef) grouped into 8 bars of common time. The following independent variable conditions were employed:

- Structure - 2 levels: structured and unstructured
- Key - 3 levels: diatonic notes from C major, G major and F major
- Preview size - 6 levels: 1, 2, 3, 4, 5 beats, and unlimited preview

As with the monophonic study, participants were required to play the sequences as fast as they could comfortably manage without sacrificing care and accuracy. Pitch and timing data were recorded for performances, and from the latter a mean tempo in beats per minute was calculated for each sequence, providing a total of 756 data points (36 trials x 21 participants) for statistical analysis - a 3-factor within subjects (structure, key and preview size), 1-factor between subjects (skill) repeated measures ANOVA.

7.1.2 Aims of the analysis

In the monophonic study the root of skill difference at the sight-reading task was found not to lie in the skilled group's processing of larger quantities of preview than the less skilled group (the specific mechanism proposed by the patterning account) but rather in their more efficient processing of effective preview levels shared by both types of reader. This indicates the modified patterning account or the perceptuo-motor account to be more appropriate explanations of skill difference, with the latter seeming more plausible considering both the lack of any significant simple effect of structure for either skill group, and the superior performance of the skilled group at small preview sizes where little unambiguous musical patterning would have been available for perception. The central focus for the current analysis, therefore, is to consider whether such conclusions continue to find support within the data of this more demanding two-part task. Research by Gilman and Underwood (2003) has indicated that skilled and less skilled readers use very similar amounts of preview during dual stave sight-reading, but the extent to which skill difference with this type of material is the result of patterning or perceptuo-motor related factors has not previously been investigated.

7.2 Results

7.2.1 Main effects

To begin with, the main effects will be considered in order to provide a general context for later, more detailed analysis. As with the monophonic study, the conditions of both skill and preview size produce highly significant effects. In relation to skill ($F\ (1, 19) = 60.62$, $p < 0.0001$), the skilled group once again performed, on average, considerably faster than the less skilled group (skilled = 155 beats per minute (SD = 46); less skilled = 89 beats per minute (SD =24). The results for preview size ($F_{7,133}\ [1, 19] = 60.62$, $p < 0.0001$) are presented in Table 7.1 and Figure 7.1. The mean tempi of the combined skill groups display a similar quadratic trend to that obtained in the monophonic study, with tempo increasing with preview, but at a declining rate until an apparent asymptote is reached.

In contrast to the monophonic study, a highly significant main effect of structure is present ($F\ (1, 19) = 88.41$, $p < 0.0001$), with structured sequences performed, on average, more quickly than unstructured ones. The mean tempo for combined skill groups on the structured level is 125 beats per minute (SD = 50) and on the unstructured level, 116 beats per minute (SD = 48). As with the monophonic study, there was no significant main effect of key ($F_{2,38}\ [1, 19] = 0.19$, $p = 0.67$).

Table 7. 1
Mean tempo (beats per minute) at levels of preview size (beats) (combined across skill, structure and key, SD in parentheses)

Preview size	1	2	3	4	5	Unlimited
Tempo	70 (15)	106 (28)	124 (39)	139 (48)	141 (51)	143 (55)

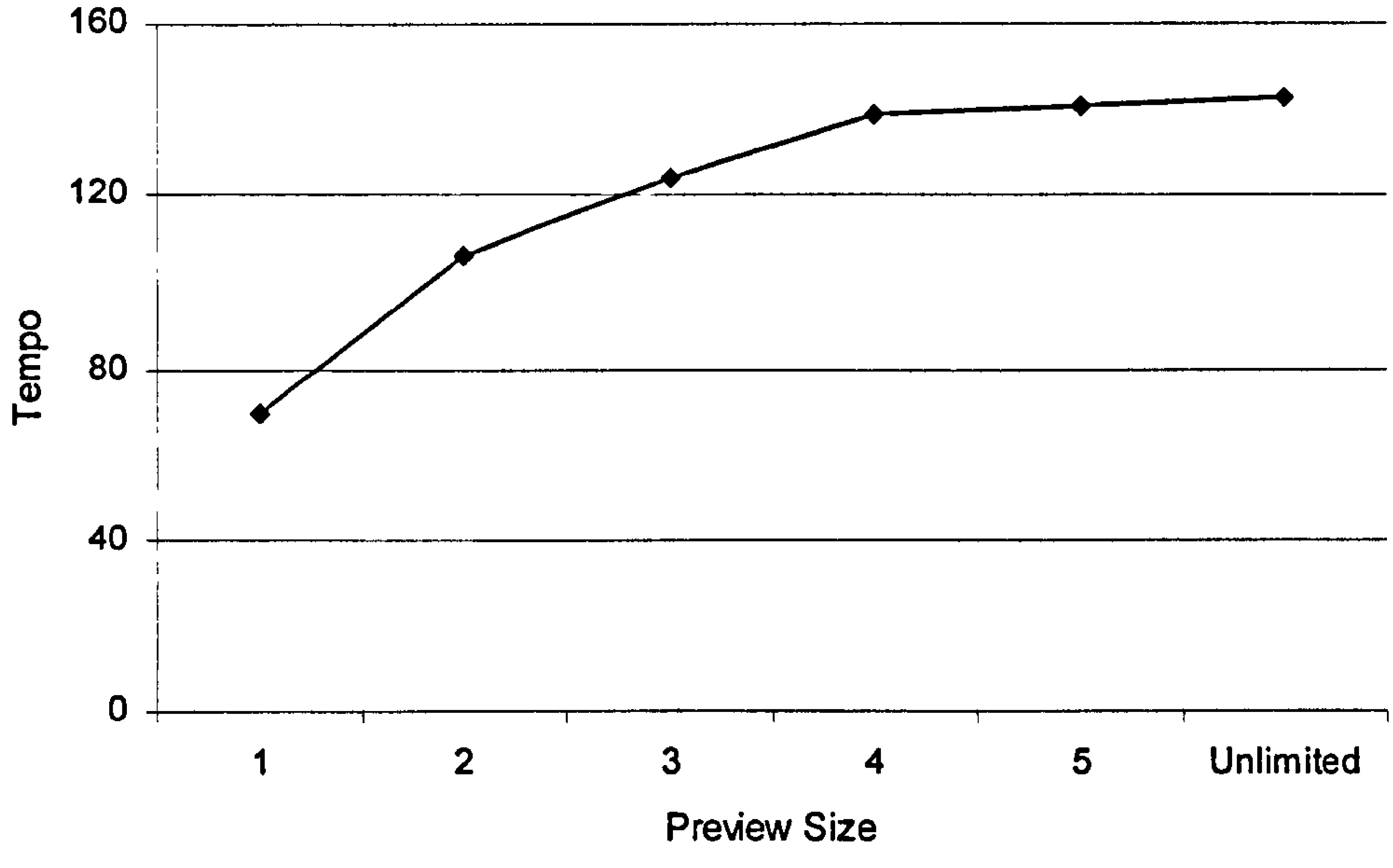


Figure 7. 1
Mean tempo (beats per minute) at levels of preview size (beats)
(combined across skill, structure and key)

7.2.2 Interactions and simple effects

As with the equivalent part of the monophonic analysis, this is divided into three sections. Firstly, a brief introduction to the patterns of preview use of the skill groups in order to set the scene. Secondly, an analysis of the role of structure within the data; and finally, in the context of these findings, a more detailed revisiting of the issues of skill and preview, particularly involving a quantification of maximum effective preview spans.

7.2.2.1 Patterns of preview use of the skill groups

The skill x preview size interaction is again highly significant ($F_{7, 133}$ [1, 19] = 36.05, $p < 0.0001$), and the preview distributions of the two groups are presented in Table 7.2 and Figure 7.2. The distribution of each subject group demonstrates the same general quadratic shape to their combined distribution in Figure 7.1, and as was the case with the monophonic study, the interaction is clearly visible. Again, the smallest tempo difference between the groups occurs at 1 beat of preview, but the skilled group now have a considerably greater performance advantage here of 23 beats per minute. The gains achieved as preview size increases are, once more, smaller for the less skilled group, and the mean tempo of these participants increasingly lags behind that of their skilled counterparts. The less skilled readers appear to reach asymptotic performance at a slightly smaller preview size to the skilled readers, but considering the highly significant main effect of structure, it would seem wise to explore the influence of this condition underlying these distributions prior to engaging in any quantification of maximum effective preview.

Table 7. 2
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key, SD in parentheses)

Preview size	1	2	3	4	5	Unlimited
Skilled	82 (11)	130 (19)	158 (22)	182 (27)	187 (31)	190 (37)
Less skilled	59 (7)	84 (16)	93 (21)	100 (22)	100 (24)	99 (23)
Skilled minus less skilled	23	46	65	82	87	91

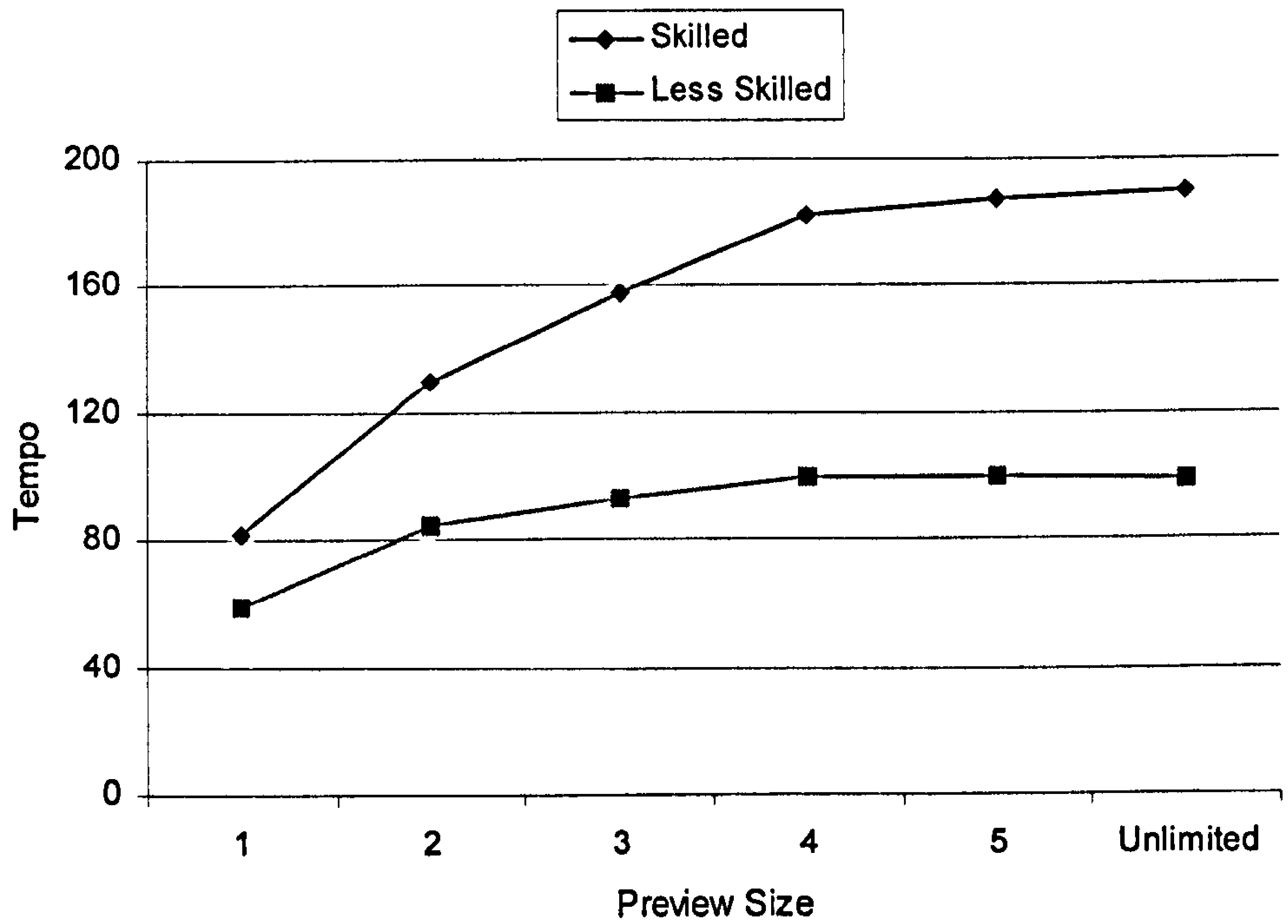


Figure 7.2
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key)

7.2.2.2 Structure and its relationship with skill and preview size

Unlike the monophonic study, there is no significant skill x structure interaction ($F(1, 19) = 0.19$, $p = 0.66$). The structure x preview interaction is at a statistically significant level ($F_{5, 95}(0.5, 8.6) = 10.73$, $p = 0.02$) and the data relating to this are displayed in Table 7.3 and Figure 7.3. The effect of structure is virtually absent

Table 7.3

Mean tempo (beats per minute) for levels of structure at levels of preview size (beats) (combined across skill and key, SD in parentheses)

Preview size	1	2	3	4	5	Unlimited
Structured	71 (14)	110 (28)	128 (38)	145 (47)	149 (51)	149 (55)
Unstructured	70 (15)	102 (29)	120 (39)	133 (48)	134 (51)	136 (54)
Structured minus unstructured	1	8	8	12	15	13

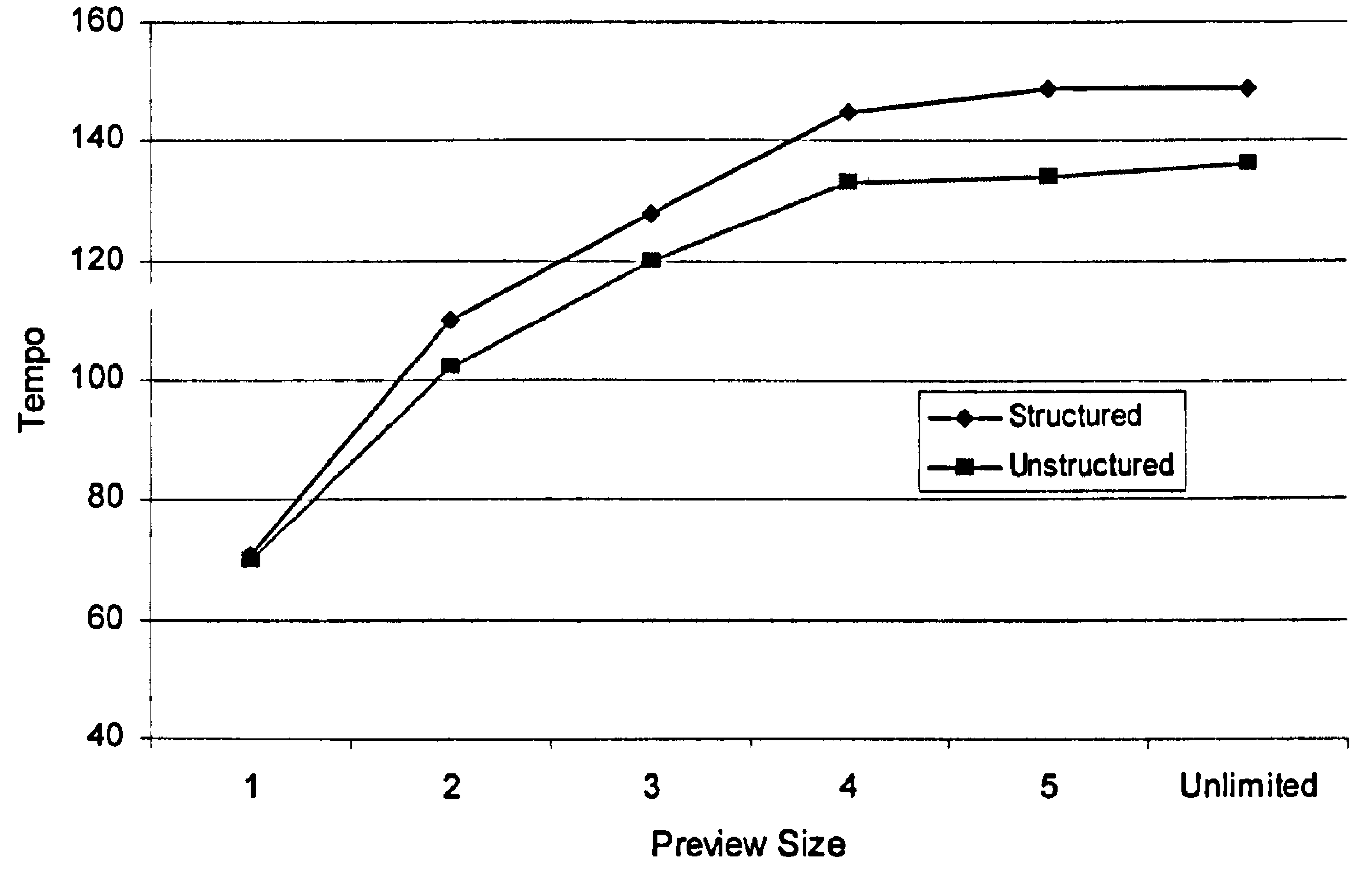


Figure 7.3

Mean tempo (beats per minute) for levels of structure at levels of preview size (beats) (combined across skill and key) at 1 beat of preview, but it then grows as preview size increases. There is not a significant 3-way interaction of skill x structure x preview size ($F_{5, 95}$ (0.5, 8.6) = 0.85, p = 0.29), suggesting that the pattern of structure shown in Figure 7.3 is not substantially different for individual skill groups. Because of this, there would not normally be a reason for investigating the individual skill group interactions. However, the faster rate of tempo increase with preview size for structured material that is at the heart of the structure x preview interaction does suggest the possibility that asymptotic performance might be reached at a larger level of preview size with this type of material, something that is not precluded by the non-significant skill x structure x preview size interaction. It would seem wise, therefore, to make separate quantifications of maximum effective preview span for structured and unstructured material.

7.2.2.3 Quantification of maximum effective preview spans

The data for the individual skill groups relating to the structure x preview interaction are presented together in Figure 7.4 as a graph, and separately in Tables 7.4 and 7.5. As was the case with the monophonic study, because of the ascending quadratic nature of the distributions, an effective way to quantify the spans is by the use of linear contrast analyses (further discussion of the rationale behind this can be found in the equivalent section of Chapter 6). The weighted comparison calculations are shown in Tables 7.6, 7.7, 7.8, and 7.9, the results indicating that the skilled readers make use of up to 4 beats of preview with both types of material, and that the less skilled readers similarly use 4 beats with structured sequences but only 3 beats with unstructured ones. Weighting 1b in Table 7.9 is only mildly non-significant, though, and therefore the latter figure should perhaps be treated with some caution i.e. it is possible that both skill groups are equivalent in terms of the amount of preview they use. Table 7.10 and Figure 7.5 present again the distributions in Table 7.4, Table 7.5 and Figure 7.4, but combine the results at the quantified maximum effective preview levels and beyond, data that will be useful in relation to later discussion. Finally, post hoc t-tests indicate that there is no significant of structure for either skill group at 1 beat of preview, the preview size that displays the smallest difference in tempo between the levels of structure (skilled: t = 0.45, p = 0.66; less skilled: t = 1.06, p = 0.30).

Table 7.4

Skilled group mean tempo (beats per minute) for levels of structure at levels of preview size (combined across key, SD in parentheses)

Preview size	1	2	3	4	5	Unlimited
Structured	82 (11)	133 (18)	161 (22)	188 (24)	194 (31)	198 (37)
Unstructured	82 (12)	126 (19)	155 (21)	176 (19)	179 (30)	182 (37)
Structured minus unstructured	0	7	6	12	15	16

Table 7.5

Less skilled group mean tempo (beats per minute) for levels of structure at levels of preview size (combined across key, SD in parentheses)

Preview size	1	2	3	4	5	Unlimited
Structured	60 (8)	88 (15)	98 (21)	105 (22)	108 (24)	105 (22)
Unstructured	58 (7)	81 (16)	88 (20)	94 (22)	92 (21)	94 (23)
Structured minus Unstructured	2	7	10	11	16	11

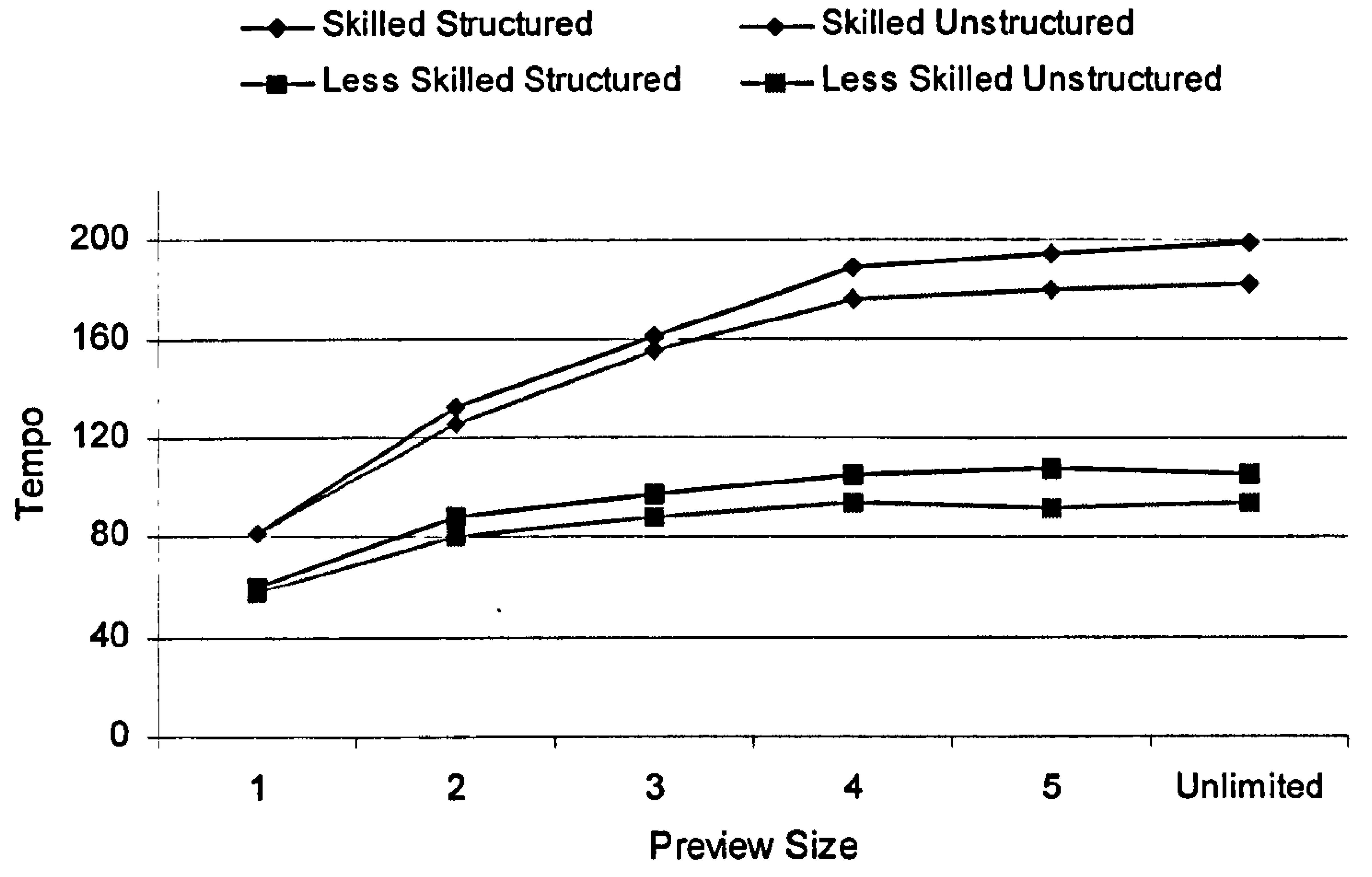


Figure 7.4
Mean tempo (beats per minute) for each skill group at levels of structure and preview size (combined key levels)

Table 7.6
Linear contrast analyses to determine maximum effective preview span: skilled group performing structured material

Preview (beats)	2	3	4	5	Unlimited	
Weighting 1a	-4	1	1	1	1	$t(45) = 10.86, p < 0.0001$
Weighting 1b	0	-3	1	1	1	$t(45) = 6.61, p < 0.0001$
Weighting 1c	0	0	-2	1	1	$t(45) = 1.59, p = 0.12$ (non. sig.)

Table 7.7
Linear contrast analyses to determine maximum effective preview span: skilled group performing unstructured material

Preview (beats)	2	3	4	5	Unlimited	
Weighting 1a	-4	1	1	1	1	$t(45) = 9.62, p < 0.0001$
Weighting 1b	0	-3	1	1	1	$t(45) = 4.79, p < 0.00001$
Weighting 1c	0	0	-2	1	1	$t(45) = 0.88, p = 0.38$ (non. sig.)

Table 7.8

Linear contrast analyses to determine maximum effective preview span: less skilled group performing structured material

Preview (beats)	2	3	4	5	Unlimited	
Weighting 1a	-4	1	1	1	1	t(50) = 6.43, p < 0.00001
Weighting 1b	0	-3	1	1	1	t(50) = 3.23, p < 0.002
Weighting 1c	0	0	-2	1	1	t(50) = 0.45, p = 0.66 (non. sig.)

Table 7.9

Linear contrast analyses to determine maximum effective preview span: less skilled group performing unstructured material

Preview (beats)	2	3	4	5	Unlimited	
Weighting 1a	-4	1	1	1	1	t(50) = 6.43, p < 0.00001
Weighting 1b	0	-3	1	1	1	t(50) = 1.95, p = 0.06 (non. sig.)
Weighting 1c	0	0	-2	1	1	t(50) = 0.52, p = 0.61 (non. sig.)

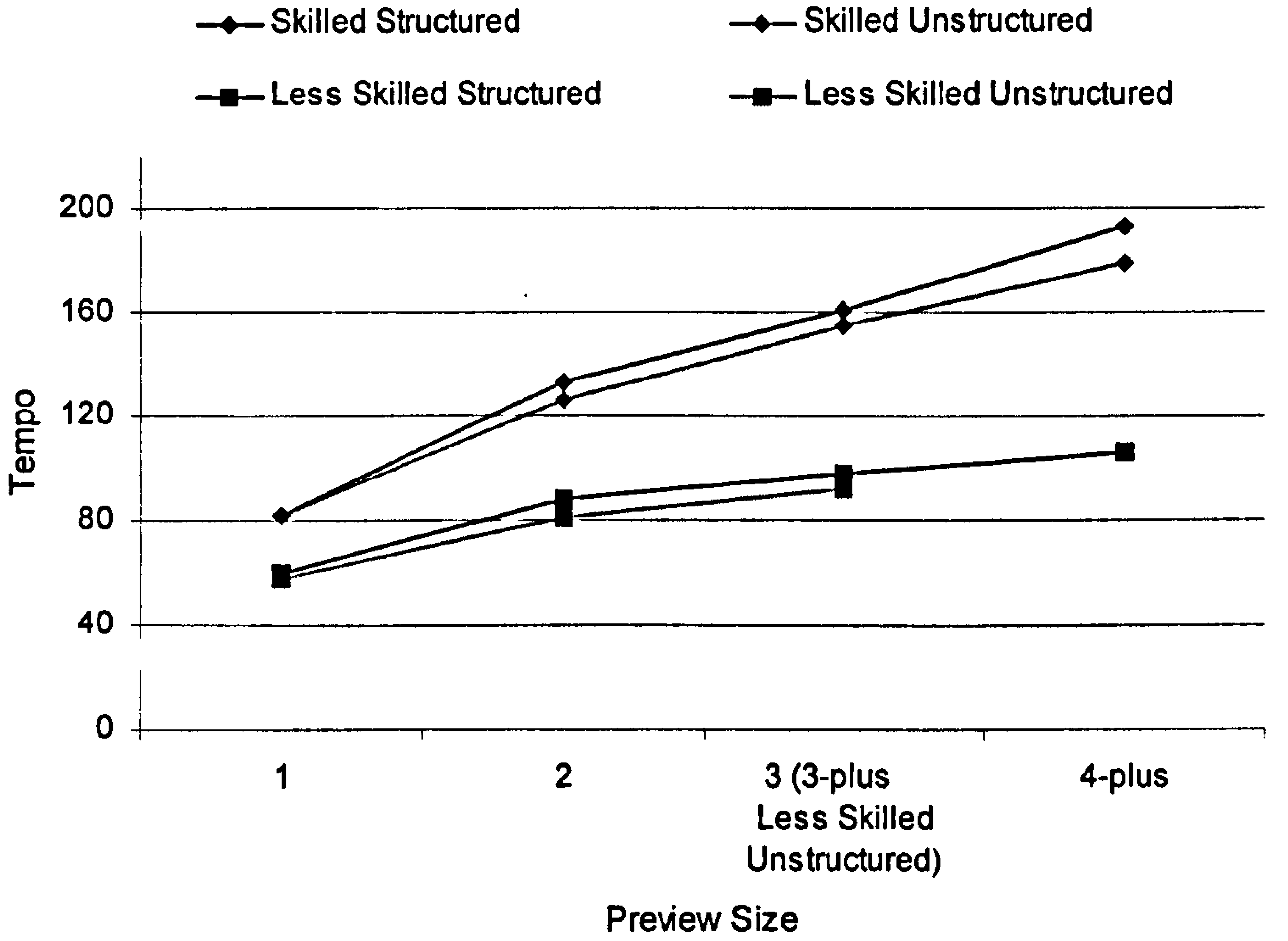


Figure 7.5

Mean tempo (beats per minute) for each skill group at levels of preview size and structure (combined across key), with results at and above maximum effective levels of preview combined

Table 7.10

Mean tempo (beats per minute) for each skill group at levels of preview size and structure (combined across key), with results at and above maximum effective levels of preview combined

Preview size	1	2	3	4-plus
Skilled - structured	82	133	161	193
Skilled – unstructured	82	126	155	179
Less skilled - structured	60	88	98	106
Less skilled – unstructured	58	81	92 (combined 3-plus preview levels)	---

7.2.3 Performance Errors

The performance error data are combined figures relating to two kinds of error: notes misplayed and notes omitted. On average, each skilled participant performed 18.7 errors throughout the entire 36 sequences (1.1% error rate), and each less skilled participant performed 49 errors (2.9% error rate).

7.2.4 Summary of Results

The skilled group performs significantly faster than the less skilled group at all levels of preview size. The significant effect of preview size produces skilled and less skilled distributions that are quadratic in shape, tempo initially increasing sharply at smaller preview sizes, followed by a tailing off of performance gains until an asymptote is reached. With the strong skill x preview size interaction, the initially small difference in tempo between the groups at 1 beat of preview (23 beats per minute) grows consistently with increasing preview availability.

There is a significant main effect of structure, but no significant skill x structure interaction, meaning that the two groups can be considered to have equivalent responses in relation to the structural distinction. A significant structure x preview size interaction is present, with greater preview leading to relatively faster performance on structured material for both groups. Less skilled participants reach asymptotic performance at 4 beats of preview with structured notation (106 beats per minute) and 3 beats with unstructured (92 beats per minute). The asymptotes for skilled readers are both achieved with 4 beats of preview (structured tempo= 193 beats per minute; unstructured tempo = 179 beats per minute).

7.3 Discussion

7.3.1 Introduction

The similarity of the maximum effective preview spans for the two skill groups with two-part notation provides confirmation of the monophonic study's questioning of the mechanism of the patterning account: again, it is efficiency of processing, not size of preview, that is found to be at the source of the superior performance of the skilled sight-readers. These results are also consistent with the findings of Gilman and Underwood (2003) who in their dual-stave study using three-part notation, discussed in detail in the literature review, demonstrated the effective preview of their skilled readers to lie between 4 and 5 beats, and that of their less skilled readers to lie between 3¾ and 4 ¾ beats. With the skilled and less skilled participants able to use very similar quantities of preview, the choice of explanation for skill difference at the task again lies with either the modified patterning account, in which sensitivity to structure may still be considered primarily responsible for driving skilled performance but perceived within a narrower notational context than proposed by the patterning account itself; or the perceptuo-motor account. This discussion begins with an assessment of which of these two accounts provides the more appropriate explanation for the difference in performance of the skill groups in this study, followed by a more detailed consideration of the results undertaken in the light of this assessment. The final section returns to an examination of the monophonic study data in the context of this current study's findings, particularly to explore again the issue of factors limiting performance.

7.3.2 An evaluation of perceptuo-motor and patterning-based accounts of the data

As with the monophonic study, gauging the appropriateness of the perceptuo-motor and modified patterning accounts as interpretations of the data requires firstly that the meaning of the effect of structure is understood. At preview sizes where effective preview can be maximised, the sight-reading of structured sequences enables both skill groups to make the same, small gain of 14 beats per minute upon their unstructured sequence performance levels (skilled/structured = 193 beats per minute, skilled/unstructured = 179 beats per minute; less skilled/structured = 106 beats per minute, less skilled/unstructured 92 beats per minute (see Table 7.10)). Perhaps not too much can be concluded from the small size of the gain. The figure very likely understates absolute sensitivity to patterning since even the unstructured sequences contained recognisable structural elements. However, despite absolute sensitivities to structure not being quantifiable from the data, it could be argued that the 14 beats per minute gains represent 'relative' sensitivities that are directly proportional to absolute levels, meaning that both types of reader can be considered to have equivalent responses to structure within the experimental materials as a whole. With the two skill groups considered similar in relation to overall sensitivity to structure in their performances, the difference in their tempi at either level of the structure condition can then be viewed as representative of the extent to which their abilities differ as a result of perceptuo-motor factors. The fact that with unstructured material the skilled group's asymptotic performance tempo is nearly twice as fast as that of the less skilled group (skilled = 179 beats per minute; less skilled = 92 beats per minute) would appear to be evidence of a considerable perceptuo-motor superiority.

The structural effect is capable of being interpreted in other ways, though. In the last chapter, a musical patterning based interpretation of the non-significant monophonic structural effect was proposed, one consistent with the modified patterning account. It was considered that the skilled group might have been sufficiently sensitive to musical structure to be able to maximise their performance on the basis of the patterning available within the unstructured sequences alone. The less skilled group, on the other hand, might have been so insensitive to structure that its greater availability within structured material was of no performance benefit to them. The significant effect of structure in the current study can be accounted for by a modification of this interpretation. Here, the effect may indicate the skilled group coming close to maximising performance on the basis of structure in the unstructured sequences, and requiring a merely partial perception of the extra structure available to them within the structured sequences to fully achieve this. With regard to the less skilled group, the structural effect may represent their entire response to the extra structure available within the structured sequences, perception of structure being the limiting factor to their performance.

It is of course possible that the recorded effect of structure may include responses that do not relate to musical patterning at all. For example, the sequences may not have been adequately controlled for performance difficulty level, resulting in the unstructured ones being harder to play than the structured, and therefore requiring slower performance. As discussed in the methodology, attempts were made to keep these two types of sequence as equivalent as possible in this regard. Certainly, the individual treble and bass stave parts were similar in nature to the monophonic sequences, and nothing that could be considered a performance difficulty effect was apparent in that study. However, such an effect clearly cannot be ruled out, meaning that some, or possibly all, of the effect of structure may be unrelated to participants' actual patterning perception

Although the modified patterning account of skill has been suggested as a possible alternative to the perceptuo-motor account, as with the monophonic study, the data at small preview sizes raise considerable doubts about its appropriateness as an explanation. If the perception of patterning were the principal motivator of skill difference, one would expect little in the way of performance variation between the skill groups at very low levels of preview (i.e. 1 or 2 beats) with unstructured material. The available notation here is highly structurally impoverished both by virtue of its unstructured nature and by the very little structural opportunity afforded by the narrow notational context. However, skilled subjects are found to perform considerably faster than less skilled subjects at these preview sizes. At 1 beat of preview the skilled group perform 41% faster than the less skilled group with unstructured material (skilled = 82 beats per minute; less skilled = 58 beats per minute (see Table 7.10). With scarcely no opportunity for the perception of structure, it is difficult to see how anything other than superior perceptuo-motor ability could account for this. By 2 beats of preview, skilled subjects are sight-reading 37% faster than the less skilled group's maximum speed with this type of material (skilled = 126 beats per minute; less skilled = 92 beats per minute). It seems unlikely that two unstructured treble stave notes and two unstructured bass stave notes could be capable of consistently communicating sufficient meaningful patterning to deliver this large performance advantage. This would require the skilled participants to be extremely sensitive to very minimal structural cues.

From the evidence as it stands, therefore, it would appear that, once more, the most reasonable explanation for the difference in the performance of the two skill groups is provided by the perceptuo-motor account. The next section of this discussion involves some further examination of the results interpreted particularly from this perceptuo-motor perspective, providing a more detailed consideration of its appropriateness.

7.3.3 Further analysis of the data from a perceptuo-motor perspective

The fact that both skill groups are able to perform 14 beats per minute faster on structured sequences than on unstructured sequences is clearly contrary to the assumption of the patterning account's proponents that less skilled readers are not as sensitive to musical structure as skilled readers. It must be emphasised, though, that this finding does not contradict prior empirical evidence, because to my knowledge, this is the first study that has actually measured structural influence on the complete task performance of less skilled sight-readers. The finding is also in line with Waters et *al.'s* perceptual sub-skill research, discussed in the literature review, which showed that in relation to pitch transformations, all their skill groups were slower with randomised than with coherent materials and that 'there was no evidence of an expertise x pitch structure interaction' (Waters *et al.*, 1997, p. 481).

The similar sensitivity to structure of the skill groups is perhaps not that surprising, considering that both consist of experienced musicians for whom the perception of structure would be an essential component of their memory organisation within rehearsed performance. Something more surprising does emerge, however, when these performance gains are considered relative to the tempi at which they were achieved. The tempo increase of 14 beats per minute represents an 8% gain for the skilled group and a 15% gain for the less skilled group, signifying that the less skilled participants take twice as much advantage (proportionately) of structure as do the less skilled participants. Although this represents a reversal of the expectations of the patterning account, such a result would seem to be entirely consistent with a perceptuo-motor based explanation of skill. For example, the superior perceptuo-motor ability of the skilled subjects may mean that they simply have less need of structure to help increase their speed of output. Less skilled subjects, on the other hand, may depend more upon the availability of familiar patterned elements within the score to compensate for their weaker perceptuo-motor functioning.

Although responses to the structural distinction may have understated participants' absolute sensitivity to musical patterning, this would only have been in relation to local structure i.e. basic structural building blocks like broken triads. The unstructured sequences were impoverished in terms of larger scale structure content like tonally coherent melodic and harmonic progression, meaning that virtually all the performance gains due to inferential and predictive mechanisms would have been limited to the performance of the structured sequences. Clearly it is impossible to differentiate between gains made from local and larger scale structure within the structured material, and with hindsight, it is clear that a further level of structure would have been useful in this study, one that consisted predominantly of local structuring, to enable some distinguishing of the two effects. However, it can at least be concluded that the variation in performance on the structure condition represents the maximum level of influence that inferential and predictive processes could have had upon performance. The fact that these processes have been responsible, at most, for only an 8% performance gain amongst the skilled subjects raises questions about the emphasis placed upon them by the patterning account, because as has previously been discussed in the literature review, Sloboda considered them relevant even to less demanding monophonic sight-reading. It may be, though, that the formal experimental conditions of my study simply hindered their use, with participants preferring to play safe and aim for note perfect performances, rather than take risks. Or perhaps the sequences were rhythmically impoverished and too general in terms of musical style for predictive purposes. A final point worth making in relation to this is that the small size of the gains achieved with structured material may have been partly due to the relatively short length of the sequences themselves. Participants may have been able to compensate for the lack of patterning in the unstructured sequences through the mustering of greater short-term concentration and effort, something that may not have been sustainable had they been required to perform more extended material.

The manner in which structure interacts with preview is similar for both skill groups and has a seemingly straightforward interpretation consistent with a perceptuo-motor account of skill. Performances on the two structure levels are virtually identical at 1 beat of preview, a result readily explained by the lack of context available for structural perception. For both skill groups at 2 and 3 beats of preview, the performance of structured material leads to larger, but still relatively small gains of 6 or 7 beats per minute. As already discussed, at asymptotic tempo levels, the performance advantage for structured sequences rises to its largest value of 14 beats per minute, again for both skill groups. The fact that the influence of structure grows with increasing preview can be accounted for in both perceptual and motor terms. With regard to the former, as the notational context increases, structural elements become more readily discernible, leading to efficiencies in perceptual uptake. In terms of the latter, larger preview enables motor activity to be more effectively planned, thereby enabling subjects to gain greater performance advantage from the more familiar, pattern based movements.

The fact that my less skilled group recorded a smaller maximum effective preview with structured material than with unstructured material is consistent with the idea that short-term memory capacity was starting to be a limiting factor to their performance of unstructured sequences i.e. more efficient, pattern-based memory storage was required to provide the extra preview necessary for sustaining faster output. The 4-beat spans recorded by the skilled group on both structure levels suggest that there was no similar constraint on memory in the performance of these participants. It must be borne in mind, though, that the difference in span for the less skilled group was only mildly significant in statistical terms, and so perhaps not too much should be read into it. This issue will be returned to in Chapter 8 where it becomes more pertinent to discussion, and where alternative explanations of variation in maximum effective preview span are considered.

7.3.4 Factors limiting performance: further consideration of the monophonic data

The maximum effective preview spans of the skilled group are two notes larger for two-part sequences than for monophonic sequences (monophonic = 6 notes, two-part = 8 notes). For the less skilled group, the equivalent increase in span is four notes for structured material (monophonic = 4 notes, two-part = 8 notes) and two notes for unstructured material (monophonic = 4 notes, two-part = 6 notes). As has been discussed earlier, proponents of the patterning account (Sloboda, 1978b) and others (Rayner & Pollatsek, 1997) have considered short-term storage capacity to be the limiting factor for monophonic performance. However, the larger maximum effective preview spans achieved by both skill groups on the two-part sequences compared to the monophonic ones suggests that that they had short-term memory storage to spare during their sight-reading of monophonic notation, pointing instead to some other perceptuo-motor factor or factors being responsible for constraining the speed of monophonic output.

These findings offer support for my proposal in the literature review that Sloboda's eye-performance spans (skilled readers = 7 notes, less skilled readers = 4 notes) might be underestimating of short-term memory capacity, particularly in relation to less skilled readers. There is also evidence within the monophonic and two-part preview study data supportive of my proposed explanation of the underestimation. The case was made that the single performance tempo employed by Sloboda may have been set too fast to enable the filling of short-term memory buffers, making the measures obtained for different participants possibly only representative of their relative speeds of encoding within the available time window for perception. To ensure that his measures reflected short-term memory capacity he should also have recorded them at progressively slower tempi, which would have provided longer time windows for encoding and so enabled an asymptote level for storage size to be determined. To test this interpretation of Sloboda's data it is clearly necessary that further eye-performance span research is carried out in the manner just described. However, my two-part study does provide an approximation to this work, the materials having been performed by both skill groups at a considerably slower tempo than in the monophonic study. The data indicate that the larger maximum effective preview spans achieved with two-part material compared to monophonic material are directly related to the longer encoding time made available by the slower tempi employed. The less skilled group recorded their maximum effective preview span of 4 notes on the monophonic study at 204 beats per minute, and their span of 8 notes with structured two-part material at a tempo of 106 beats per minute. Therefore, at just over half their maximum monophonic tempo, and so in the context of twice the amount of time available for perception, the less skilled group were able to encode information effective to performance from twice the quantity of notes. My skilled readers' monophonic maximum effective preview span of 6 notes relates to a tempo of 299 beats per minute, and their span of 8 notes on the two-part structured sequences to 193 beats per minute. This latter figure represents approximately two-thirds of their monophonic tempo, providing them with half as much time again for encoding, and therefore the opportunity to increase their memory storage with information from an extra 3 notes - making 9 notes in total.

The above interpretation rests upon the assumption that there is no difference in how monophonic and two-part material is stored in memory. This is not necessarily the case, though. For example, it is possible that individual parts are stored in discrete areas, each with individual capacity limits. More research into memory for monophonic notation is needed, therefore, to test the validity of the assumption. Also, although slower performance tempi allowing the encoding of larger quantities of notes would seem to account very effectively for the magnitude of the increases in maximum effective preview span achieved by the skill groups with two-part material, other explanations are possible. For example, as discussed in the literature review, in other domains it has been found that increasing cognitive load can lead to a diminishing of perceptual span (Henderson & Ferreira, 1990). It is therefore conceivable that the task demand of maximum tempo monophonic performance may be greater than that of the considerably slower maximum tempo associated with two-part performance, and that the extra task load of the former may be acting to limit short-term memory capacity. Clearly, further dedicated research is required to enable final conclusions to be drawn about the origins of the differences between monophonic and dual stave perceptual measures. However, the mere fact that significant differences have been recorded suggests that the proponents of the patterning account may have been a little presumptuous in assuming that perceptual data relating to the sight-reading of simple, single line notation, would be sufficiently representative of sight-reading as a whole to provide a reliable foundation for a general theory. With the findings of the current study confirming Gilman and Underwood's (2003) dual stave preview data, discussed earlier, evidence is mounting that the patterning account is based upon an understatement of the preview capabilities of less skilled sight-readers. Bearing the dangers of presumption in mind, perhaps not too much should be read into the similarity of the maximum effective preview spans of the two skill groups in this study prior to the analysis of the data relating to the even more complex four-part sequences, which will be undertaken in the next chapter.

7.3.5 Summary of discussion

As with the monophonic study, there is no evidence to support the proposal of the patterning account that the different abilities of the skilled and the less skilled participants are primarily the result of the skilled using larger quantities of preview. Whilst the results at asymptotic tempi are capable of supporting the modified patterning account, the considerably superior performance of skilled subjects at one and two beats of preview, where little or no structure is available, makes a primarily pattern-based explanation of skill difference implausible. An alternative explanation of the result is that the skilled readers simply have faster basic perceptuo-motor abilities than the less skilled readers, and that for both groups sensitivity to musical patterning is responsible for only relatively small variation in performance tempo.

The responses of the two groups to the structural distinction are virtually identical in terms of absolute tempo, with performance gains due to structure's influence increasing with preview size, presumably caused by the increasing availability of structure within the score, and capable of both perceptual and motor explanation.

The fact that both skill groups effectively perceived a greater quantity of preview within the two-part sequences than within the monophonic sequences, suggests that short-term memory was not a limiting factor to monophonic performance for either skill group. The greater level of preview used by skilled readers on the monophonic task can be accounted for simply in terms of skilled readers being able to process notes from the score more quickly. These findings suggest that measures of monophonic preview may not typically be representative of levels of preview used with other types of notation, something that questions the validity of theories based largely upon monophonic empirical evidence. In relation to two-part material, there is possible evidence that the performance of less skilled subjects may be beginning to be limited by short-term memory storage availability.

Chapter 8

Four-part controlled preview study

8 Four-part controlled preview study

8.1 Introduction

8.1.1 Résumé of methodology

The 10 skilled and 11 less-skilled participants performed at sight, with hands together, 30 different four-part sequences (dual stave). All sequences were made up of 32 semibreve chords (2 notes in the treble stave, 2 in the bass stave) grouped into 8 bars, each bar containing 4 beats. The following independent variable conditions were employed:

- Structure - 2 levels: structured and unstructured
- Key - 3 levels: diatonic notes from C major, G major and F major
- Preview size - 5 levels: 1, 2, 3, 4 beats, and unlimited preview

As with the monophonic and two-part studies, participants were required to play the sequences as quickly as they could comfortably manage without sacrificing care and accuracy. Pitch and timing data were recorded for the performances, and the grouped statistical analysis carried out using the mean tempo for each of the 630 trials (36 trials x 21 participants) - a 3-factor within subjects (structure, key and preview size), 1-factor between subjects (skill) repeated measures ANOVA.

8.1.2 Aims of the analysis

In the previous two studies, the principal source of the difference between the two skill groups' performances has been found not to be the skilled readers' use of greater preview than the less skilled readers, but rather their faster processing of equivalent levels of preview. Whilst it remains possible that sensitivity to musical structure may be primarily responsible for the observed skill difference (the modified patterning account), the superior performance of skilled participants with unstructured material at very small preview sizes suggests a fundamentally perceptuo-motor account to be more plausible. Therefore, a priority for this analysis is to examine the extent to which the data continue to show evidence of the above trends.

There are two other points particularly requiring attention. Firstly, interpreted in the context of the perceptuo-motor account, both skill groups showed similar, fairly limited responses to the structural distinction in their two-part performances. Will this pattern of response be replicated with four-part notation? Certainly, with the heavier task demand of four-part material it might be expected that participants overall will make greater use of structure in order to maximise their performances. Secondly, in the two-part study there was possible evidence of the less skilled readers' output starting to be limited by short-term memory storage availability. If this was in fact the case, one would expect to see further evidence in this current study, given the greater complexity of the materials involved.

8.2 Results

8.2.1 Main effects

The general picture provided by the main effects will be considered prior to more detailed statistical investigation. All four main effects of skill, preview, structure and key are statistically significant, the former three effects particularly so. With regard to skill ($F (1, 19) = 41.06, p < 0.0001$), the mean tempo of the skilled group on structured sequences, at combined levels of preview size and key, is 105 beats per minute (SD = 26), and that of the less skilled group is 59 beats per minute (SD = 17). Data relating to preview size ($F_{4, 76} [1, 19] = 106.65, p < 0.0001$) are presented in Table 8.1. They show participants performing at what appears to be two levels of tempo, a slower rate at 1 beat of preview followed by a step change up to a considerably faster, relatively constant rate at 2 beats and beyond.

Table 8.1
Mean tempo (beats per minute) at levels of preview size (beats) (combined across skill, structure and key, SD in parentheses)

Preview size	1	2	3	4	Unlimited
Tempo	60 (15)	82 (29)	87 (33)	86 (34)	88 (35)

Turning to structure ($F(1, 19) = 109.99, p < 0.0001$), at combined skill, preview size and key levels, the mean tempo for structured tests is 86 beats per minute (SD = 33), and for unstructured, 75 beats per minute (SD = 30). Finally, the significant main effect of key ($F_{2, 38}[1, 19] = 13.28, p = 0.002$) indicates that I have been unsuccessful in my attempts to control the ease of performance of sequences across the key levels. However, the resulting differences in tempi are only very small: C = 81 beats per minute (SD = 32), F = 82 beats per minute (SD = 32) and G = 80 beats per minute (SD = 31). In view of this, it would be most unlikely that the effect of key is having any distorting influence upon the analysis that would affect the general conclusions. In retrospect, though, it would clearly have been preferable to formally pilot the tests first of all to ensure the greatest degree of equivalence.

8.2.2 Interactions and simple effects

This part of the analysis is again divided up into three sections. First, there is a brief introduction to the patterns of preview use of the skill groups. Second, an investigation into the role of structure within the data, and third, a return to issues of preview and skill, considered in more detail in the context of the analysed structural effects.

8.2.2.1 Patterns of preview use of the two skill groups

The skill x preview interaction is once again significant ($F(4, 76) = 5.12$, $p = 0.04$), and the relevant data are presented in Table 8.2 and Figure 8.1. These individual distributions of skill as a function of preview size are similar in shape to the combined distribution already presented in Table 8.1. The interaction between them is clearly visible between 1 and 3 beats of preview. Starting off with a performance advantage of 22 beats per minute at 1 beat of preview, the skilled group then makes a considerably larger tempo gain than the less skilled group from 1 to 2 beats, the former increasing speed by 34 beats per minute, the latter only by 10 beats per minute. At 2 beats, the less skilled group appear to have reached peak performance, but the skilled group continue to make further, although much diminished, gains up to 3 beats. Further investigation of preview together with statistically based quantifications of the maximum effective preview spans will be undertaken after some understanding has been gained into how these distributions are influenced by structure.

Table 8.2
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key, SD in parentheses)

Preview size	1	2	3	4	Unlimited
Skilled	72 (11)	106 (19)	114 (21)	115 (24)	117 (24)
Less skilled	50 (9)	60 (17)	63 (18)	61 (17)	63 (19)
Skilled minus less skilled	22	46	51	54	54

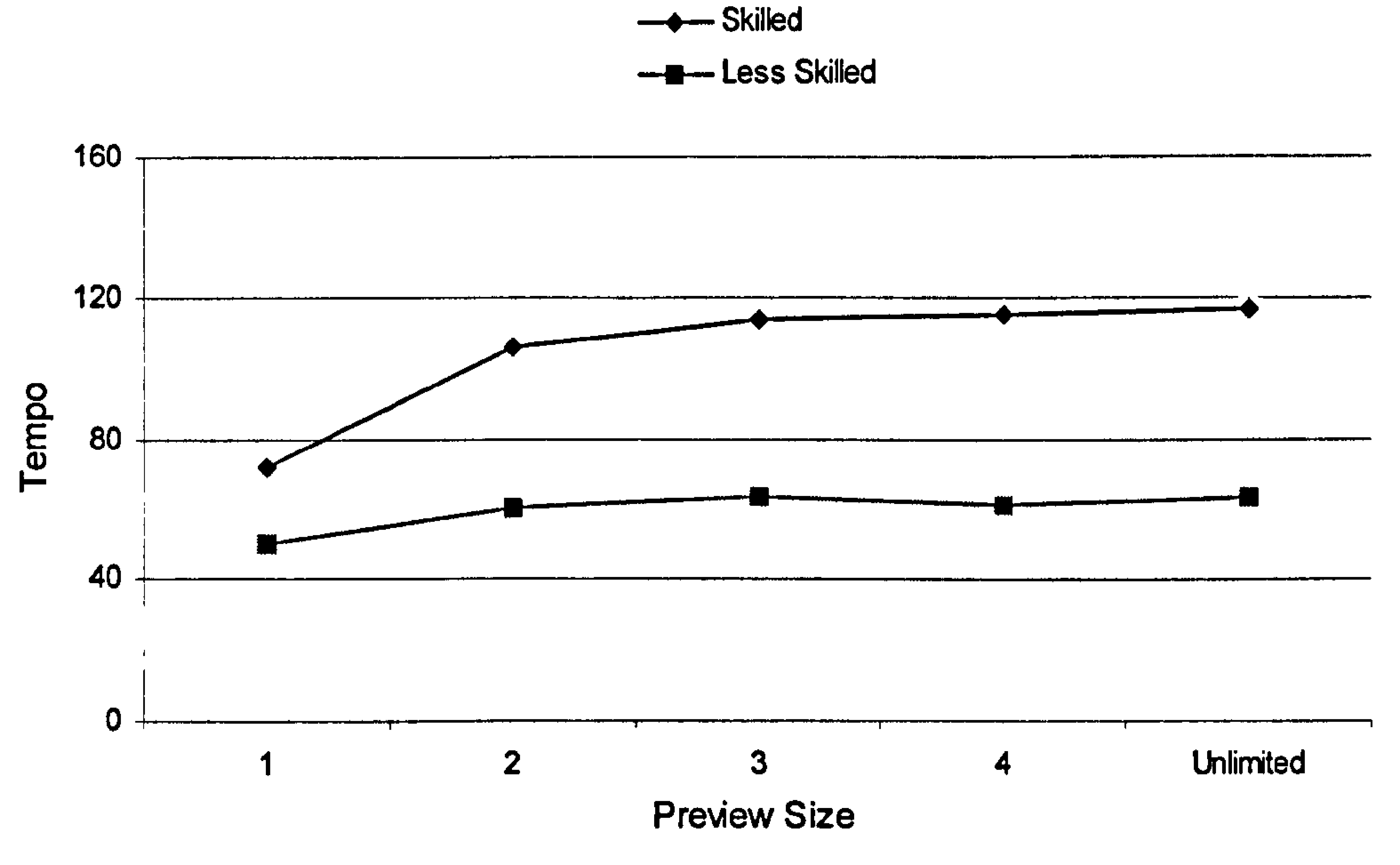


Figure 8.1
Mean tempo (beats per minute) for each skill group at levels of preview size (beats) (combined across structure and key)

8.2.2.2 Structure and its relationship with skill and preview size

The pattern of interactions with relation to structure is equivalent to that found with two-part notation. There is no significant skill x structure interaction present ($F (1, 19) = 0.4, p = 0.53$), the analysis therefore again failing to distinguish between the skill groups in terms of response to the structural distinction. The interaction of preview size x structure is also significant ($F_{4, 76}$ (0.44, 8.44) = 12.47, p = 0.02), and the data relating to this are displayed in Table 8.3 and Figure 8.2. The distributions show a small performance benefit with structured material at 1 beat of preview, becoming considerably larger and reasonably consistent at 2 beats of preview and beyond.

Table 8.3
Mean tempo (beats per minute) for levels of structure at levels of preview size (beats) (combined across skill and key, SD in parentheses)

Preview size	1	2	3	4	Unlimited
Structured	62 (15)	88 (29)	95 (33)	93 (35)	95 (36)
Unstructured	58 (15)	76 (28)	80 (31)	81 (32)	82 (33)
Structured minus unstructured	4	12	15	12	13

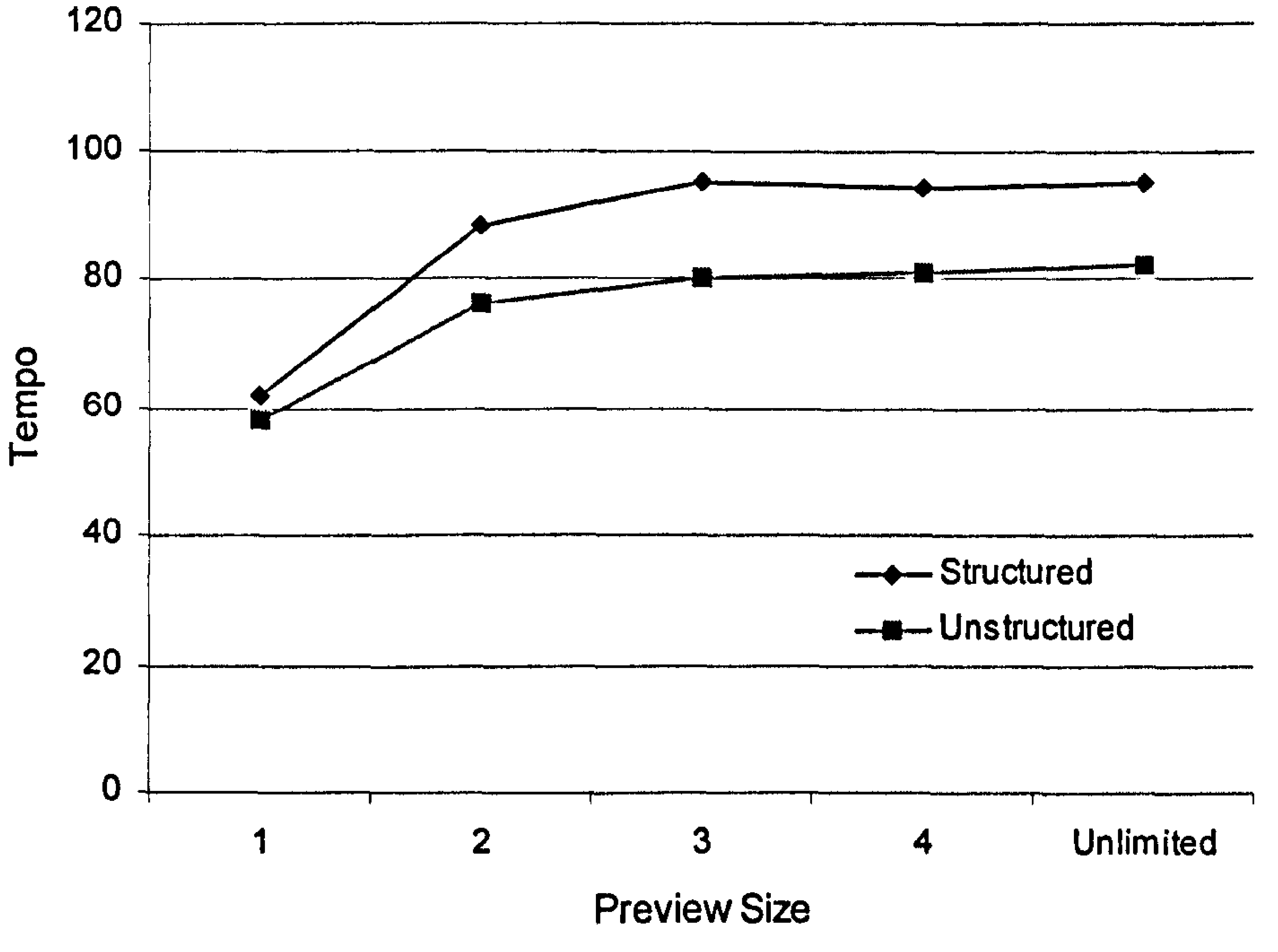


Figure 8.2
Mean tempo (beats per minute) for levels of structure at levels of preview size (beats) (combined across skill and key)

As with the two-part study, because there is no significant 3-way skill x structure x preview interaction ($F_{4, 76}$(0.44, 8.44) = 0.67, p = 0.32), an investigation of the preview x structure interactions of the individual skill groups would not normally be justified. However, the significant structure x preview size interaction again suggests the possibility that asymptotic performance may be reached at a larger preview size with structured material than with unstructured material, an effect that is not necessarily ruled out by a non-significant skill x structure x preview interaction. It would seem wise, therefore, to once more make separate quantifications of maximum effective preview span for structured and unstructured materials.

8.2.2.3 Determination of maximum effective preview spans

The results for the individual skill groups relating to the structure x preview interaction are displayed in Table 8.4 and Table 8.5 respectively, and also together in graph form in Figure 8.3. Once again, the maximum effective preview spans of the two skill groups performing structured and unstructured material have been determined using linear contrast analyses. The four weighted comparison calculations are presented Tables 8.6, 8.7, 8.8 and 8.9. The analyses indicate that span does not vary with structure for four-part notation, skilled readers achieving spans of 3 beats and less skilled readers 2 beats on both levels of the structure condition. Table 8.10 and Figure 8.4 present again the distributions in Table 8.4, Table 8.5 and Figure 8.3, but combine the results at the calculated maximum effective preview levels and beyond. These data will be useful in relation to later discussion. Finally, post hoc t-tests indicate that the effect of structure is significant for each skill group at 1 beat of preview, the preview size exhibiting the smallest difference in tempo between the levels of structure (skilled: $t = 6.33$, $p < 0.0001$; less skilled: $t = 6.03$, $p < 0.0001$).

Table 8.4

Skilled group mean tempo (beats per minute) for levels of structure at levels of preview size (combined across key, SD in parentheses)

Preview size	1	2	3	4	Unlimited
Structured	74 (12)	112 (19)	123 (20)	121 (24)	124 (26)
Unstructured	70 (10)	101 (17)	107 (19)	108 (22)	110 (21)
Structured minus unstructured	4	11	16	13	14

Table 8.5

Less skilled group mean tempo (beats per minute) for levels of structure at levels of preview size (combined across key,SD in parentheses)

Preview size	1	2	3	4	Unlimited
Structured	51 (9)	66 (18)	69 (19)	66 (18)	68 (19)
Unstructured	47 (8)	54 (14)	56 (15)	56 (14)	56 (17)
Structured minus unstructured	4	12	13	10	12

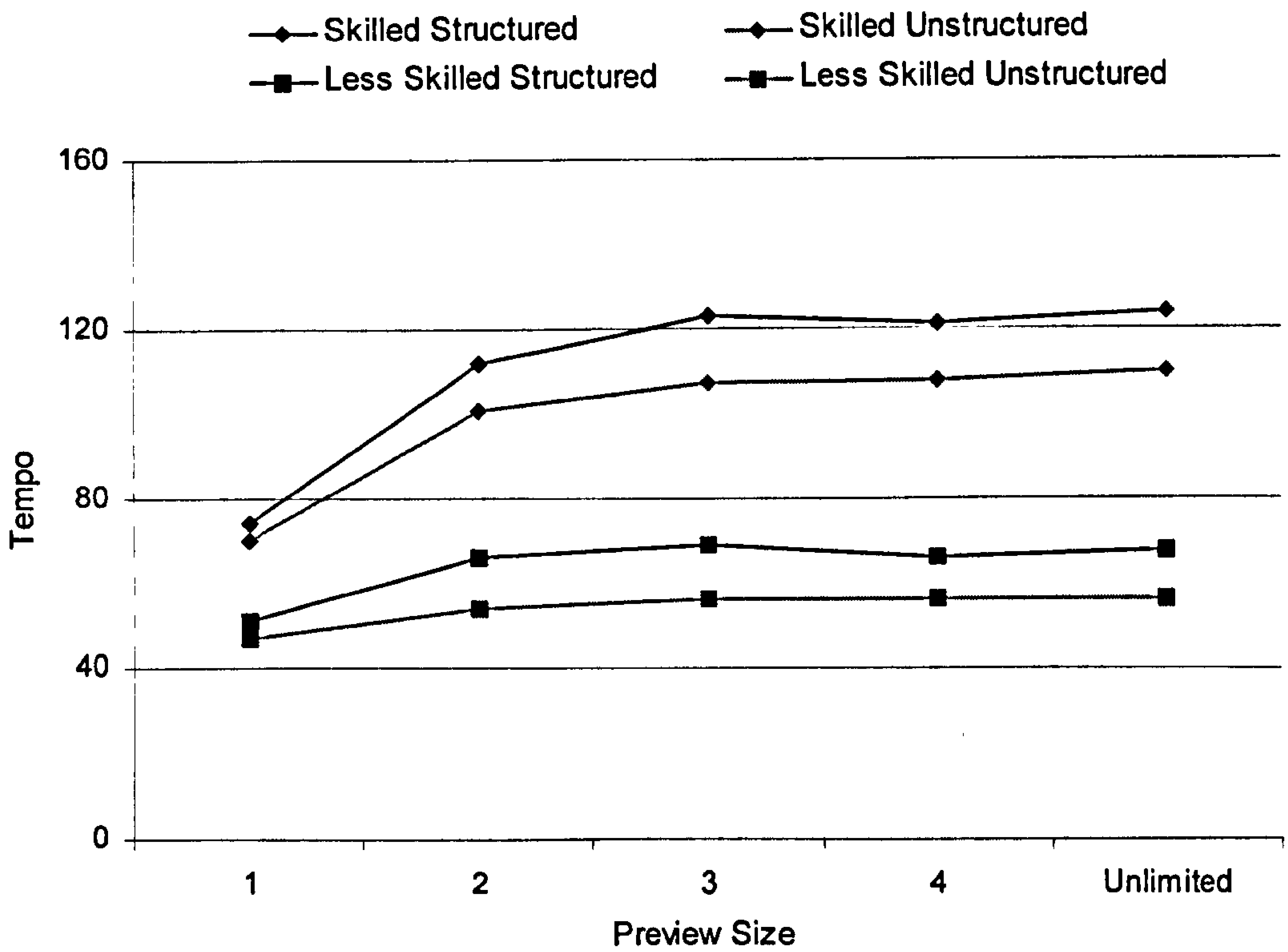


Figure 8.3

Mean tempo (beats per minute) for each skill group at levels of structure and preview size (combined across key)

Table 8.6

Linear contrast analyses to determine maximum effective preview span: skilled group performing structured material

Preview (beats)	1	2	3	4	Unlimited	
Weighting 1a	-4	1	1	1	1	t(36) = 17.02, p < 0.0001
Weighting 1b	0	-3	1	1	1	t(36) = 3.65, p = 0.0008
Weighting 1c	0	0	-2	1	1	t(36) = 0.04, p = 0.97 (non. sig.)

Table 8.7

Linear contrast analyses to determine maximum effective preview span: skilled group performing unstructured material.

Preview (beats)	1	2	3	4	Unlimited	
Weighting 1a	-4	1	1	1	1	t(36) = 16.58, p < 0.0001
Weighting 1b	0	-3	1	1	1	t(36) = 3.35, p = 0.002
Weighting 1c	0	0	-2	1	1	t(36) = 0.86, p = 0.40 (non. sig.)

Table 8.8

Linear contrast analyses to determine maximum effective preview span: less skilled group performing structured material

Preview (beats)	1	2	3	4	Unlimited	
Weighting 1a	-4	1	1	1	1	t(40) = 9.07, p < 0.0001
Weighting 1b	0	-3	1	1	1	t(40) = 0.70, p = 0.49 (non. sig.)

Table 8.9

Linear contrast analyses to determine maximum effective preview span: less skilled group performing unstructured material

Preview (beats)	1	2	3	4	Unlimited	
Weighting 1a	-4	1	1	1	1	t(40) = 6.12, p < 0.0001
Weighting 1b	0	-3	1	1	1	t(40) = 1.25, p = 0.22 (non. sig.)

Table 8.10

Mean tempo (beats per minute) for each skill group at levels of preview size and structure (combined across key), with results at and above maximum effective levels of preview combined

Preview size	1	2 (2-plus for Less skilled)	3-plus
Skilled structured	74	112	122
Skilled non-structured	70	101	108
Less skilled structured	51	67	----
Less skilled non-structured	47	56	----

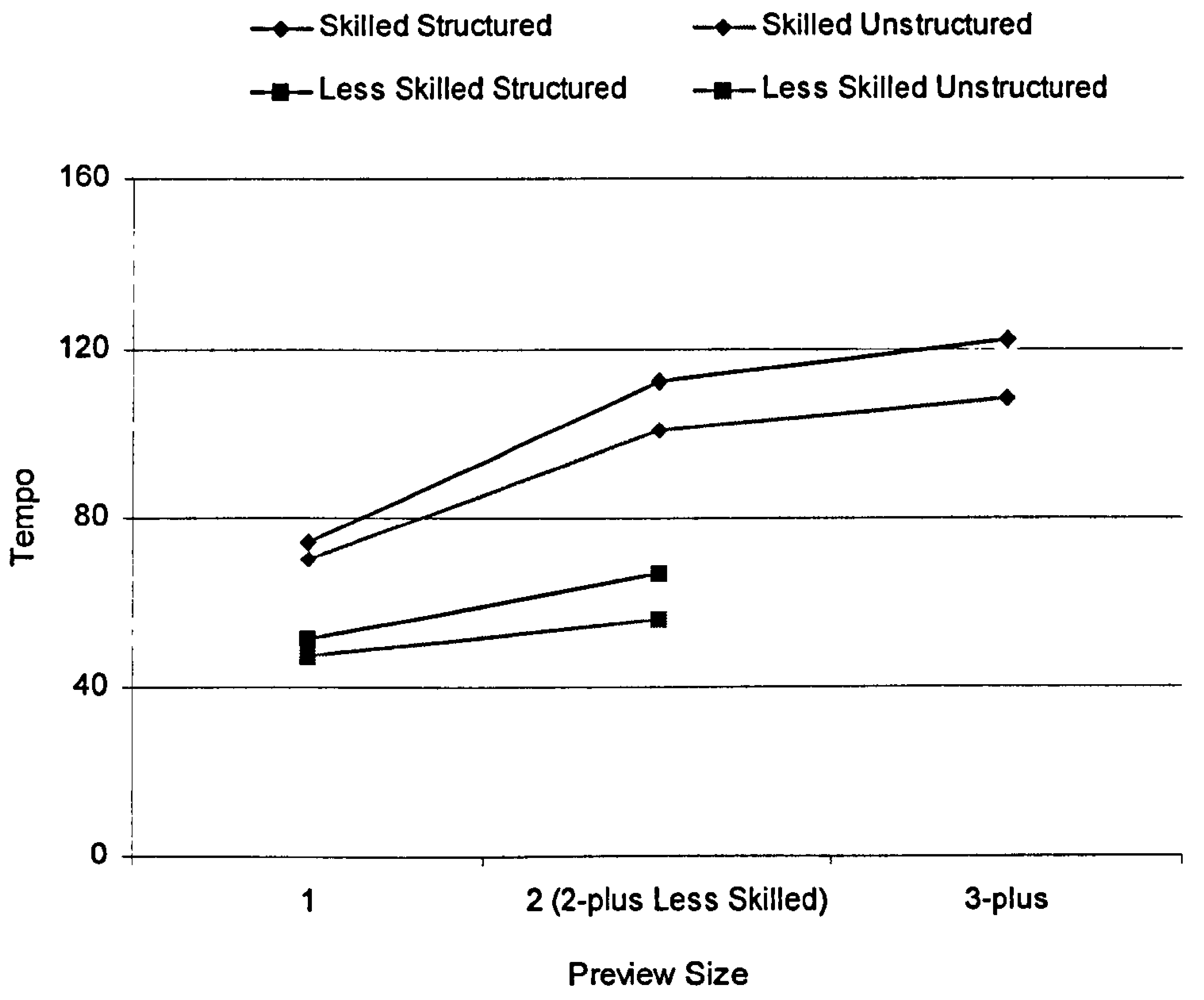


Figure 8.4
Mean tempo (beats per minute) for each skill group at levels of preview size and structure (combined across key), with results at and above maximum effective levels of preview combined

8.2.3 Performance Errors

The performance error data are combined figures relating to two kinds of error: notes misplayed and notes omitted. On average, each skilled participant performed 33.5 errors throughout the entire 30 sequences (1% error rate), and each less skilled participant performed 101.6 errors (3% error rate).

8.2.4 Summary of results

The skilled group performs significantly faster than the less skilled group at all levels of preview. The shape of the skilled reader preview distribution is similar to that obtained in the previous studies, tempo initially increasing sharply with preview size, followed by a tailing off of performance gains to an asymptote, in this case at 3 beats. The less skilled readers, on the other hand, only show a significant tempo difference between the 1 and 2-beat preview levels. There is a strong skill x preview size interaction, the difference in mean tempo of the skill groups increasing from 23 beats per minute at 1 beat of preview to 55 beats per minute at asymptotic levels of performance, with structured material.

There is a significant main effect of structure across all levels of preview, but not a significant skill x structure interaction, meaning that the two groups can be considered to have equivalent responses in relation to the structural distinction. There is also a significant preview x structure interaction, with greater preview leading to relatively faster performance on structured material for both groups. The skilled group reaches asymptotic performance at the 3-beat preview size for both structured and unstructured material (structured tempo = 122 beats per minute; unstructured tempo = 108 beats per minute). The less skilled group achieves asymptotic performance at the 2-beat preview size, also for both levels of structure (structured tempo = 67 beats per minute; unstructured tempo = 56 beats per minute)

8.3 Discussion

8.3.1 Introduction

As with the monophonic and two-part studies, the data are again not supportive of the patterning account's proposal that the source of the superior performance of skilled sight-readers lies in their use of greater preview. Although with this four-part material the skilled group have maximum effective preview spans that are one beat larger than the less skilled group, their tempo at only 1 beat of preview is faster than their less skilled counterparts' asymptotic performance (1-beat preview skilled/structured = 74 beats per minute, 1-beat preview skilled/unstructured = 70 beats per minute; 2-beat preview less skilled/structured = 67 beats per minute, 2-beat preview less skilled/unstructured = 56 beats per minute). By two beats of preview, the skilled readers demonstrate a considerable performance advantage - 55 beats per minute faster than the less skilled readers with structured material (skilled = 112 beats per minute, less skilled = 67 beats per minute). The use of the third beat of preview only increases the mean skilled group tempo for structured material by a further 10 beats per minute to 122 beats per minute, a gain of 9%. Therefore, the evidence once more points to a choice between the modified patterning account, in which the root of skilled readers' ability is considered to lie primarily in the perception of patterning within levels of preview effective to both skill-types, and the perceptuo-motor account, in which skill difference is viewed as being principally due to skilled readers' more efficient perceptuo-motor processing.

The rest of this chapter follows a similar course to the previous two chapters. First of all, consideration is given to whether it is the modified patterning account or the perceptuo-motor account that provide the more appropriate explanation of the data, with further discussion of the results then undertaken in the light of the conclusions drawn. The issue of factors limiting performance is then returned to. Finally, the asymptotic data from all three experiments are examined together.

8.3.2 An evaluation of perceptuo-motor and patterning based accounts of the data

The magnitude of the responses of the two skill groups to the structural distinction is similar to that of the two-part study. At 2 beats of preview both skilled and less skilled readers perform on average 11 beats per minute faster on structured than on unstructured material (skilled/structured = 112 beats per minute, skilled/unstructured = 101 beats per minute; less skilled/structured = 67 beats per minute, less skilled/unstructured = 56 beats per minute). At 3 beats of preview the tempo difference increases to 14 beats per minute for skilled readers (skilled/structured = 122 beats per minute, skilled/unstructured = 108 beats per minute).

These responses to structure again have a number of possible interpretations, lending themselves to both perceptuo-motor and patterning based accounts of overall skill difference at the task. As in the two-part study, it can be argued that the effect of structure is a partial measure of sensitivity to musical patterning, proportional to the absolute sensitivity of each group. According to this interpretation, the two groups' similar responses to the structural distinction are therefore representative of their similar sensitivity to structure within the experimental materials as a whole, pointing to the difference in their performance on each of the levels of structure being the result of perceptuo-motor factors. However, it is again possible that the apparently similar responses to structure have a more complex underlying pattern of causation. Once more, it could be that the skilled participants are considerably more sensitive to structure within the unstructured sequences than the less skilled participants, and perceive sufficient patterning within them to achieve close to maximum performance tempi, with further gains on structured material achieved by only a partial processing of the extra structure available. The less skilled participants on the other hand may be highly insensitive to structure within the unstructured materials, with their small performance gains on structured material demonstrating the full extent of their sensitivity to the available patterning. Another issue that must continue to be borne in mind is that other factors may be involved in the structural effect. For example, it may be the case that the two levels of structure were not sufficiently controlled for performance difficulty.

Whilst at asymptotic levels of performance the data might be considered capable of supporting both the modified patterning account and the perceptuo-motor account, as with the other studies, the former account would seem less plausible when data at smaller levels of preview are taken into consideration. At 1 beat of preview, skilled participants perform 49% more quickly than less skilled participants with unstructured notation (skilled = 70 beats per minute, less-skilled = 47 beats per minute). Although the four-part nature of the material provides more opportunity for structural perception within a single beat than was the case in the previous studies, it is difficult to envisage how sensitivity to patterning could be primarily responsible for the large performance advantage, in the absence of consistent and obvious structural content. Overall, therefore, the perceptuo-motor account would again seem to be the most credible explanation for skill difference at the task, and will provide the context within which the data will be considered for the remainder of the chapter.

8.3.3 Further discussion of the effect of structure

When the magnitude of the structural effect is considered in terms of performance tempo, confirmation is found for the rather surprising finding in the two-part study that the less skilled group take greater advantage of structure in their performances than the skilled group. At asymptotic levels of output, the skilled readers performed 13% faster with structured material than with unstructured material, and the less skilled readers 20% faster. Again, a possible explanation for this is that the superior perceptuo-motor abilities of the skilled readers enabled them to achieve closer to maximum performance levels before they required the assistance of structural cues.

The equivalent increases for the two-part study were 8% and 15% respectively. This greater relative sensitivity of both groups to structure with four-part material compared to two-part material is in line with the expectation, outlined at the beginning of the chapter, that dependency upon structure might be related to task demand. However, the larger gains might also be attributed to the fact that the greater number of notes in the four-part materials provided more opportunity for structure to be perceived.

A further similarity between the four-part and the two-part results is in relation to the structure x preview interaction: in both cases, the effect of structure grows with increasing preview size. At 1-beat preview with the four-part data, both skill groups have a statistically significant advantage of 4 beats per minute with structured material. The two-part material afforded no significant advantage here, the faster performance on the four-part material consistent with the point made earlier that chords of 4 notes are capable of communicating a greater amount of structural information than chords of 2 notes. With 2 beats of preview, both groups increase their gains with structured material to 11 beats per minute, with the third beat of preview enabling skilled readers to perform 14 beats per minute faster, as already discussed. Like the equivalent two-part study interaction, this pattern of interaction is entirely consistent with the perceptuo-motor account, and interpretable in both perceptual and motor terms, larger preview allowing a wider context for meaningful structure to be perceived, and facilitating better motor planning that enables pattern-based movement efficiencies to be more effectively expressed.

8.3.4 Factors limiting performance

In the two-part study, the fact that the less skilled participants used one beat less preview on unstructured than on structured sequences was considered possible evidence that short-term memory storage was a limiting factor to the performance of the former materials i.e. the larger preview necessary to achieve faster output was dependent upon achieving pattern-based memory storage efficiencies. In the four-part study, the maximum effective preview span for the less skilled group, measured in beats, is the same for both structured and unstructured material. However, because each beat consists of four notes, there is no guarantee that participants were actually making use of the complete second beat of preview; performances at either level of structure may have involved only some of the notes within the beat being processed. The fact that in their use of the second beat of preview these readers did make smaller tempo gains with unstructured than with structured material is consistent with the idea of fewer notes being processed with the former. Skilled readers also make smaller gains with unstructured material on their third beat, indicating perhaps that they, too, are reaching the limits of short-term memory storage with this type of material. It is also possible, though, that in both cases the smaller gains simply reflect a slower response to the perception of the entire beat, and so it is clear that further work needs to be undertaken to clarify the situation.

There is other evidence consistent with the idea that the four-part performance of less skilled readers is limited by short-term memory. These readers use a maximum of 8 notes of preview in this current study, and despite considerably slower performance tempi than on the two-part study (affording them time to encode further notes into memory) this figure represents no gain upon their two-part preview attainments, suggesting that short-term memory storage has reached an asymptote. This may indicate, then, that memory capacity is acting as a limiting factor to their performance with both two-part and four-part materials. Skilled readers on the other hand may use up to 12 notes of preview with four-part material, representing further gains upon their two-part performance. Therefore, short-term memory storage would appear not to be limiting to their two-part performance. To gain insight into whether it is limiting in relation to four-part material, skilled readers need to be tested with more complex notation, to see if they are capable of larger storage capacities. Overall, the evidence is consistent with the skilled readers having larger raw short-term memory storage capacity than the less skilled readers. However, it is of course possible that what appear to be short-term memory limitations in the case of less skilled readers, may in fact relate to the influence of some other factor. For example, their smaller effective preview may be restricted due to difficulties with motor output. To achieve further insight into this, the individual study of perceptual and motor related sub-skills is necessary.

The fact that the Sloboda's skilled readers could not sustain their normal 7-note eye-performance spans when they performed disrupted material suggests that this figure is close to the limits of short-term memory storage for music reading (Sloboda, 1974, 1977). This raises the question of how my skilled participants were apparently able to consistently store information about more notes than this, even with unstructured notation. To begin with, it needs to be pointed out that no research appears to have been carried out into memory for dual stave notation and so it is possible that larger quantities of raw storage are readily achievable. Also, as discussed in Chapter 6, my measures, unlike those of Sloboda, were recorded with the score constantly available to participants, enabling a continual refreshing of memory that may have allowed a greater quantity of storage to be maintained. However, findings from research into touch-typing and text reading suggest two other ways in which the memory load upon my participants may have been less than it first appears. In relation to touch-typing, it was discussed in the literature review that one or two characters ahead of the point of performance are typically irrevocably committed to performance i.e. they have passed into execution buffers and therefore place little continuing demand upon short-term memory resources. Work in text reading has found that material at the forward extent of effective preview, 'can be partially processed prior to fixation' (Gilman & Underwood, 2003, p. 202). In the context of music reading this might mean, for example, that basic information like the contour of oncoming notation might be effective to performance, enabling a general priming of movement before the notes are identified in detail. These findings suggest, therefore, that it may only be the notes in the central area of effective preview that require detailed identification and storage in memory. Further research is clearly necessary to discover the extent to which these mechanisms are relevant to music sight-reading.

8.3.5 Analysis of asymptotic performance on all preview studies

A comparison of the two skill groups' mean asymptotic tempi on each of the three studies reveals some interesting points. The data, combined across structure and key conditions, are displayed in Table 8.11 and Figure 8.5. First of all, skilled participants' performance on the two-part study is quite similar to that of less skilled participants on the monophonic study (skilled/two-part = 186 beats per minute; less skilled/monophonic = 204 beats per minute). This suggests that the different tasks are of roughly equivalent task demand for the two groups. The same is the case for skilled performance on the four-part sequences and less skilled performance on the two-part sequences (skilled/four-part = 115 beats per minute; less skilled/two-part = 99 beats per minute). On average, therefore, the skilled participants would seem to experience roughly the equivalent task demand of less skilled participants when they sight-read material that contains twice the number of parts.

Table 8.11
Mean asymptotic tempo (beats per minute) for each skill group on the three controlled preview studies (combined across structure and key)

	Monophonic	Two-part	Four-part
Skilled	299	186	115
Less skilled	204	99	62

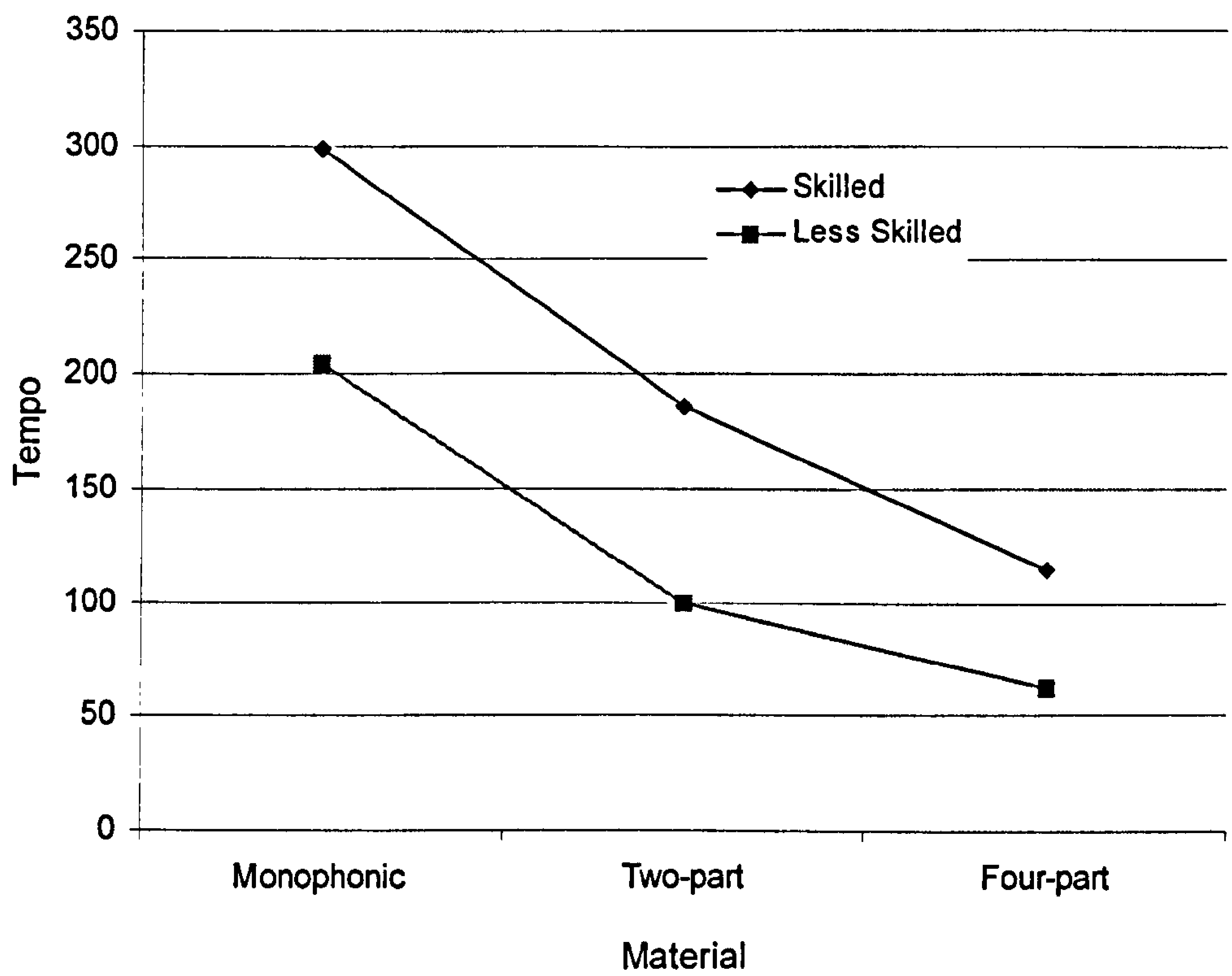


Figure 8.5
Mean asymptotic tempo (beats per minute) for each skill group on the three controlled preview studies (combined across structure and key)

Further insights can be gained by considering the data in terms of notes per minute rather than beats per minute. The transformations are presented in Figure 8.6 and Table 8.12 (the two-part results from Table 8.11 having been multiplied by two, and the four-part results by four). Skilled participants are 54% more efficient at individual note processing with four-part material than with monophonic material (monophonic = 299 notes per minute; four-part = 460 notes per minute). The equivalent increase in efficiency for less skilled participants is only 22% (monophonic = 204 notes per minute; four-part = 248 notes per minute). The skilled readers make identical

Table 8.12
Mean asymptotic note performance rate (notes per minute) for each skill group on the three controlled preview studies (combined across structure and key)

	Monophonic	Two-part	Four-part
Skilled	299	372	460
Less Skilled	204	198	248

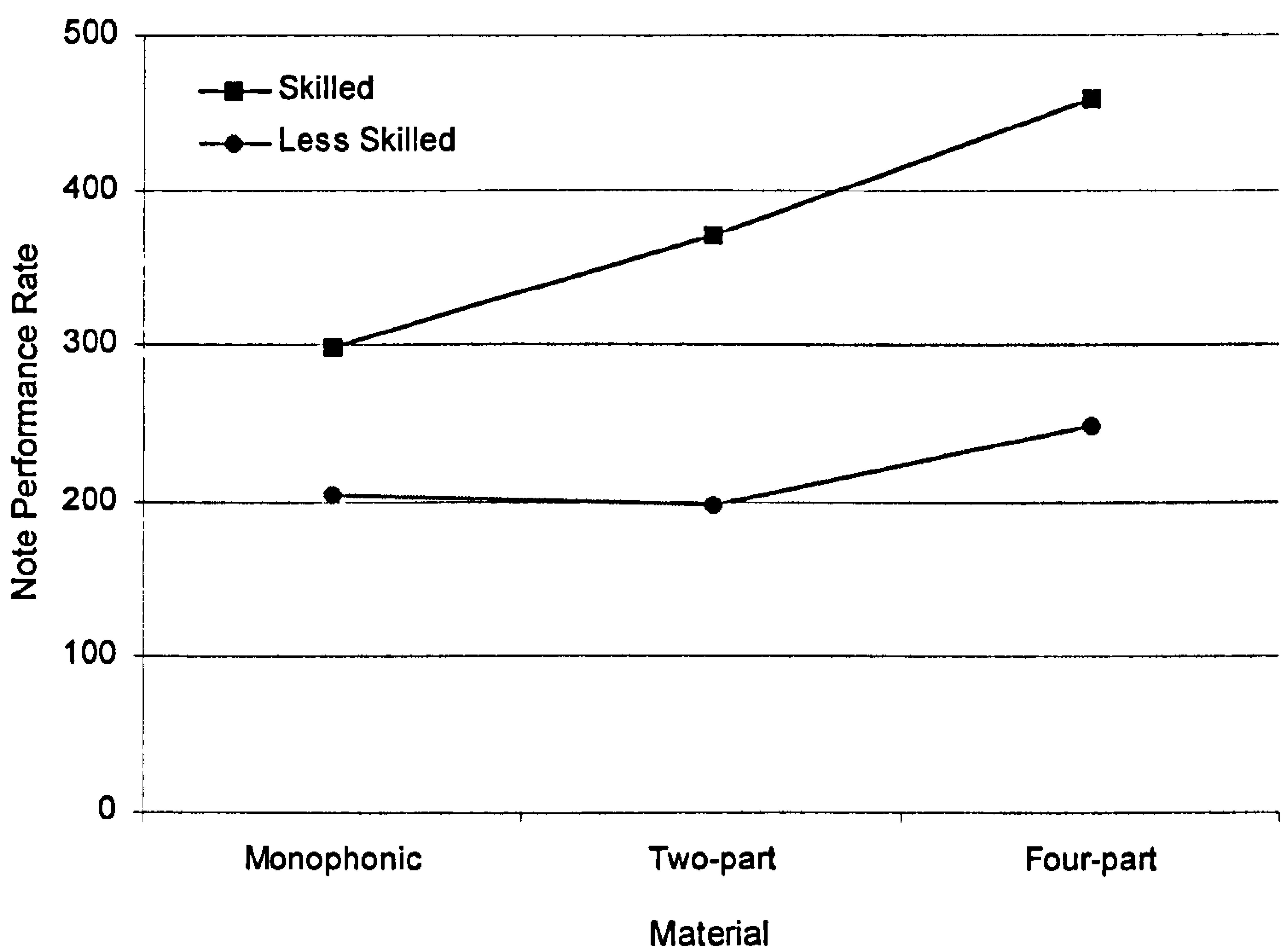


Figure 8.6
Mean asymptotic note performance rate (notes per minute) for each skill group on the three controlled preview studies (combined across structure and key)

percentage performance gains of 24% from monophonic to two-part material, and from two-part to four-part material. The less skilled readers on the other hand become 3% less efficient in their rate of note processing with two-part sequences compared with monophonic ones (monophonic = 204 notes per minute, two-part = 198 notes per minute), but then make a gain of 25% from two-part to four-part notation. Although similar in percentage terms to the equivalent tempo increase of the skilled readers, this latter figure represents a considerably smaller absolute gain (skilled increase = 88 beats per minute; less skilled increase = 50 beats per minute). The fact that the individual parts of the two-part material are similar in nature to the monophonic sequences, and so not easier in terms of required movement, indicates that the gain in rate of note processing from the latter to the former material by the skilled readers requires a perceptual processing explanation. The slower two-part tempi may have provided them with more time within which to demonstrate spare capacity in the rate of individual note processing; or perhaps these readers used note interval relationships between the staves to increase efficiency of processing. Although the less skilled readers did not achieve any increase in the rate of note processing from monophonic to two-part material, their four-part performance clearly indicates that they had the potential for further processing rate gains. A number of factors might account for these further processing efficiency gains, and a consideration of them provides some further possible insight into why the less skilled group failed to make similar gains between monophonic and two-part performance.

First of all, it could be that perception becomes more efficient because of the greater 'density' of notes with four-part material, meaning that more information can be gathered from individual fixations. It might also be that there is more opportunity for participants to make use of strategies based upon intervallic relationships within the larger chords, with less dependence upon merely individual note identification. Motor related explanations are also possible. First of all, the planning of fingering may have been easier with four-part material, their block chord nature necessarily acting as a considerable constraint upon fingering choice within individual parts, making movement planning easier. Also, although there were more notes to play within these sequences, the typically smaller within-part intervallic movement, particular in relation to the middle two parts, may have made many of the individual movements easier to execute. Another possibility is that the sight-reading of block chords involves less overlapping of perceptual and motor processing i.e. the emphasis is more upon preparing groups of notes for simultaneous rather than staggered performance. This last explanation might also be relevant to explaining why less skilled readers did not make processing efficiency gains between monophonic and two-part material. It is possible that they found the extent of parallel processing involved with two-part material (i.e. 4 beats of preview, like monophonic material, but with two separate parts to be processed) particularly hampering to performance. One explanation for this might be a lack of automation in perceptual and motor mechanisms leading to interference between them. A final point to mention is that there is no evidence to indicate whether the less skilled group are less efficient with dual stave material compared to the skilled group because of a specific hands-together performance effect, or because of mechanisms associated with their inferior single-hand performance.

8.3.6 Summary of discussion

The current study confirms the findings of the monophonic and two-part studies. Again, the root of skill difference is clearly found not to lie in skilled readers' use of greater preview than less skilled readers, but in their more efficient use of equivalent quantities of it. The evidence once more suggests that the skilled group owe their ability to possessing faster perceptuo-motor skills than the less-skilled group, with both groups displaying limited sensitivity to musical structure in their performances. However, a primarily patterning-based explanation of skill cannot be completely discounted by the data. Similar to the two-part study, responses of both groups to the structural distinction are virtually identical in absolute terms, but when the different performance speeds of the groups are taken into account, the less-skilled readers' responses demonstrate relatively greater gains with structured material. The preview x structure interaction for each skill group is also similar in shape and magnitude, the effect of structure again increasing with preview size. Once more, this can be accounted for by the greater availability of structure at larger preview levels, but whether the causative factors are perceptual or motor remains indeterminate. Finally, the larger maximum effective preview span of the skilled group may indicate that they have larger raw short-term memory capacity than the less skilled group in relation to the transcription of musical notation; other interpretations of this evidence are possible, though.

Chapter 9

Analysis of individual participant data for the controlled preview studies

9 Analysis of individual participant data for the preview studies

9.1 Introduction

Significant doubts have been raised in the controlled preview studies about a primary role for musical patterning in determining sight-reading ability, the evidence instead pointing to skill difference being mainly the result of variation in the speed of perceptuo-motor processing in relation to the task. The preview studies are not capable of identifying any particular perceptual or motor bottlenecks responsible for limiting performance. For this, the complete task results need to be considered in the context of data relating to the perceptual and motor skills of participants, something that is the focus of Chapters 10, 11 and 12. Before doing this, though, it is important that the most complete picture possible about full task performance is built up from the preview research data. In particular, further understanding is required about the relationship between participants' performances on the three studies. A comparison of grouped data was undertaken in the last chapter, but it is possible that some individuals may have shown substantial variation from the clear trends demonstrated there, indicative of other factors being involved in determining their abilities. Correlation analysis would clearly be useful to explore this (and indeed is fundamental to the work undertaken in this chapter) but this technique is limited in its ability to discern potentially important behaviour that relates to perhaps only a small minority of group members. The fact that variation in skill may be accounted for in a variety of ways by a perceptuo-motor account suggests that there is certainly the potential for such sub-patterns within the distributions. To investigate these more idiosyncratic elements of the data, a more detailed, descriptive investigation is therefore required in addition to correlation analysis.

There are some potentially significant problems associated with this approach, however. Firstly, there is a validity issue: it is possible that the responses of some individuals may not be true reflections of their sight-reading skill, but express other influences, for example, tiredness or varying levels concentration. Secondly, individual data are less reliable than grouped data, involving considerably fewer replications. There is therefore the risk that some patterns discerned during analysis might simply be the effects of experimental error. Little can be done *post hoc* with regard to the former concern, but to attempt to minimise the risk of the latter, the investigation has been limited to asymptotic performance and issues of a non-structural nature, giving the individual data studied the largest replication base possible i.e. several preview levels, and all structure and key levels. Even though any conclusions drawn must necessarily be tentative, this does not negate the value of the approach, which is primarily seen in its potential to generate ideas towards which further, more robust, research can be directed.

A number of research questions are central to the analysis:

- Are some participants relatively slower or faster than their fellow group members in relation to the performance of some types of material?
- Are there clear trends visible for some participants across all types of material i.e. a continuous relative performance decline from monophonic to four-part material?
- Are there patterns of behaviour shared by a number of participants?
- Do members of the skilled group exhibit a skilled level of performance with all three types of material? Do the less-skilled participants consistently perform at less skilled levels, or is there evidence for some that their overall skill is constrained by difficulty with a particular category of material?
- Is a high level of performance on all three types of material necessary to support a high level of overall skilled sight-reading?

To facilitate the analysis, three mean values representing asymptotic performance on monophonic, two-part and four-part material (based on the results in Chapters 6, 7 and 8 respectively) have been calculated for each participant, as follows:

Monophonic material

Skilled: mean of performance at 6, 7 and unlimited beats of preview
Less skilled: mean of performance from 4 to unlimited beats of preview

Two-part material

Skilled and less skilled: mean of performance at 4, 5 and unlimited beats of preview

Four-part material

Skilled: mean of performance at 3, 4 and unlimited beats of preview

Less skilled: mean of performance from 2 to unlimited beats of preview

As mentioned earlier, to calculate each value, individual performance data have been combined across structure and key. This would seem justified considering the limited size of the structural effect, its similarity for the two skill groups and the minimal influence of key within the data. Each participant has been given a unique identifier, for example, S5 or L3, 'S' or 'L' denoting membership of the skilled or less skilled group respectively, and the number representing their within-group ranking at monophonic asymptotic performance. Thus, S1 is the fastest member of the skilled group here, and L11 the slowest member of the less skilled group. The only analytical statistical measure that has been used in this investigation is the Spearman rank correlation coefficient (r_s) using a 5% significance level. This correlation method has been chosen because it is more dependable when comparing distributions with considerably different ranges and subject to outlying values.

The analysis is divided into four sections:

- An introduction to the general trends in, and relationships between the distributions
- A detailed analysis of the less skilled participant data
- A detailed analysis of the skilled participant data
- A conclusion and summary of findings

9.2 Analysis

9.2.1 General trends and relationships

Table 9.1 and Figure 9.1 present the mean asymptotic tempo data of individual participants for monophonic, two-part and four-part material. In Figure 9.1 the data points for each type of material have been joined to make them easier to distinguish.

Table 9.1
Mean asymptotic tempo (beats per minute) for all individual participants performing monophonic, two-part and four-part notation (combined across structure and key), in descending order of monophonic tempo value (within group)

Skilled Group				Less Skilled Group			
Participant	**Mono**	**Two-part**	**Four-part**	**Participant**	**Mono**	**Two-part**	**Four-part**
S1	374	233	118	**L1**	280	134	66
S2	338	223	167	**L2**	242	114	77
S3	333	162	98	**L3**	215	89	50
S4	321	218	110	**L4**	211	105	69
S5	309	182	124	**L5**	208	106	65
S6	272	175	125	**L6**	208	102	77
S7	270	180	112	**L7**	200	86	44
S8	268	148	93	**L8**	199	122	90
S9	255	162	104	**L9**	188	108	64
S10	247	179	102	**L10**	155	66	41
				L11	153	67	36

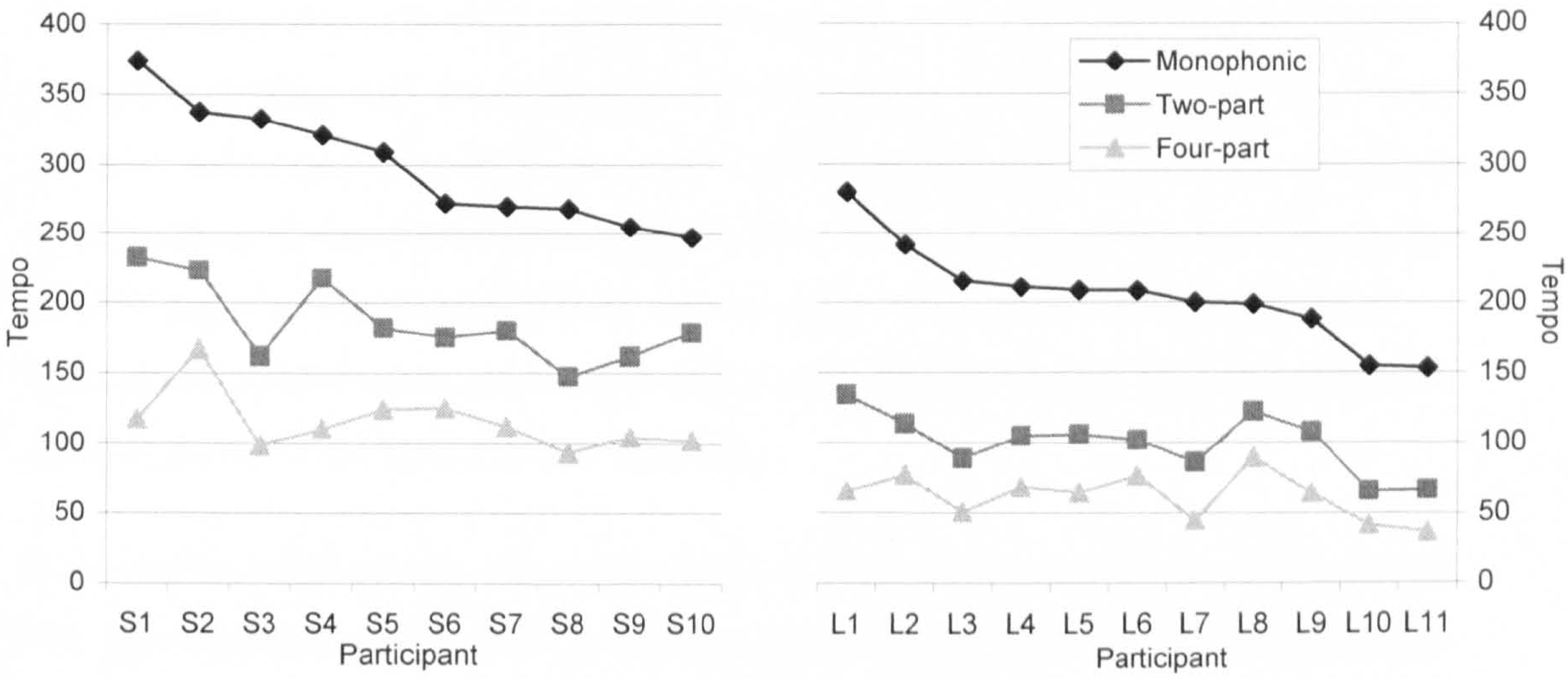


Figure 9.1
Mean asymptotic tempo (beats per minute) for individual participants performing monophonic, two-part and four-part notation (combined across structure and key), in descending order of monophonic tempo value (within-group)

The monophonic tempo data follow a gradually descending trend from S1 to L11, with an increase to L1 at the beginning of the less skilled group before the trend is resumed. L1's score is well within the skilled range, lying between those of S5 and

S6, and L2 performs only 5 beats per minute (2%) slower than S10, the slowest skilled participant with this material. After L2, there is a central group of seven less skilled readers (L3 to L9) performing considerably more slowly, all within a range of 27 beats per minute, followed by the two slowest performing participants, L10 and L11. The trend in two-part and four-part distributions is also gradually descending, but Figure 9.1 also indicates what appears to be significant variation in the performances of a number of individuals in both skill groups across the three types of material, something that will be explored shortly. There is no overlap in scores between the skill groups with either two-part or four-part notation. With two-part notation, the fastest less skilled reader (L1) performs 14 beats per minute (10%) slower than the slowest skilled reader (S8) indicating a clear hiatus between the skilled and less skilled groups. With four-part material the equivalent difference is only 3 beats per minute, in this case between S8 and L8, but after L8 there is a marked decline in less skilled performance, her closest rivals, L2 and L6, scoring 13 beats per minute (14%) more slowly.

These results confirm that the participants were accurate in the self-categorisation of their sight-reading skill. Even though 3 participants, L1, L2 and L6, scored at, or very close to, skilled group levels on one or two of the studies, their ability here was clearly insufficient to define their overall sense of skill. There are clearly too few participants to draw any final conclusion, but the results suggest that an overall skilled level of sight-reading requires the attaining of at least the minimum speed recorded within the skilled group, for all three types of material.

For combined skill groups, there are strong correlations between the monophonic and two-part distributions ($r_s = 0.87$, $p < 0.0001$), the two-part and four-part distributions ($r_s = 0.93$, $p < 0.0001$), and the monophonic and four-part distributions ($r_s = 0.82$, $p < 0.0001$). The equivalent within-group correlations are less consistent, though:

Skilled group

Monophonic and Two-part:	$r = 0.65$,	$p = 0.04$
Two-part and Four-part:	$r = 0.69$,	$p = 0.03$
Monophonic and Four-part:	$r = 0.48$,	$p = 0.16$ (non.sig.)

Less skilled group

Monophonic and Two-part:	$r = 0.55$,	$p = 0.08$ (non.sig.)
Two-part and Four-part:	$r = 0.76$,	$p = 0.007$
Monophonic and Four-part:	$r = 0.51$,	$p = 0.11$ (non.sig.)

Considering that monophonic and four-part sight-reading represent the greatest contrast of task complexity studied, it is not surprising to discover that for both skill groups these are the most weakly correlated pair of results. For the skilled group, neither of the two incremental steps of notational complexity leads individually to a significantly different pattern of performance. However, for less skilled participants, there is a significant correlation between two-part and four-part material but not between monophonic and two-part material. The fact that neither the monophonic/two-part nor the monophonic/four-part correlations are significant for the less skilled group may indicate a dual stave effect within their data i.e. perhaps some less skilled participants found sight-reading hands-together material more difficult than others. Also, considering that the tempo difference for the less skilled readers between two-part and four-part material is considerably smaller than that between monophonic and two-part material, their strongly significant two-part/four-part correlation could be explained by their having only limited spare performance capacity beyond two-part performance, meaning that there is limited opportunity for further variance to be expressed.

I will now turn to a more detailed consideration of individual data. As discussed earlier, the aim is to discover whether any participants exhibit important variation in performance with different notational complexities, and if so to attempt to discern any patterning in their behaviour. Detecting variation in the data is a simple process, but attributing the source of it to particular individual or individuals can be more problematic, the reason being that the small sample sizes can sometimes make it difficult to determine which participants form the main trend and which the deviation. In view of this, the analysis is principally concerned with identifying the more substantial sources of variation between distributions, and conclusions must be generally considered as only tentative. The data of less skilled readers will be examined to begin with, followed by those of skilled readers. For each skill group, the analysis begins with a comparison of monophonic and two-part results, followed by one for two-part and four-part results.

9.2.2 Less skilled participants

Considering monophonic and two-part performance first, Table 9.2 and Figure 9.2 present the mean tempo data of individual participants for these two types of material at maximum effective preview. Table 9.2 also includes the ranking of the scores for

Table 9.2

Mean asymptotic tempo (beats per minute) for individual less skilled participants performing monophonic and two-part notation, in descending order of monophonic tempo value (combined across structure and key, rank order of two-part performance in parentheses)

	Monophonic	Two-part (rank)	Difference
L1	280	134 (1)	146
L2	242	114 (3)	128
L3	215	89 (8)	126
L4	211	105 (6)	106
L5	208	106 (5)	103
L6	208	102 (7)	106
L7	200	86 (9)	114
L8	199	122 (2)	77
L9	188	108 (4)	81
L10	155	66 (11)	89
L11	153	67 (10)	87

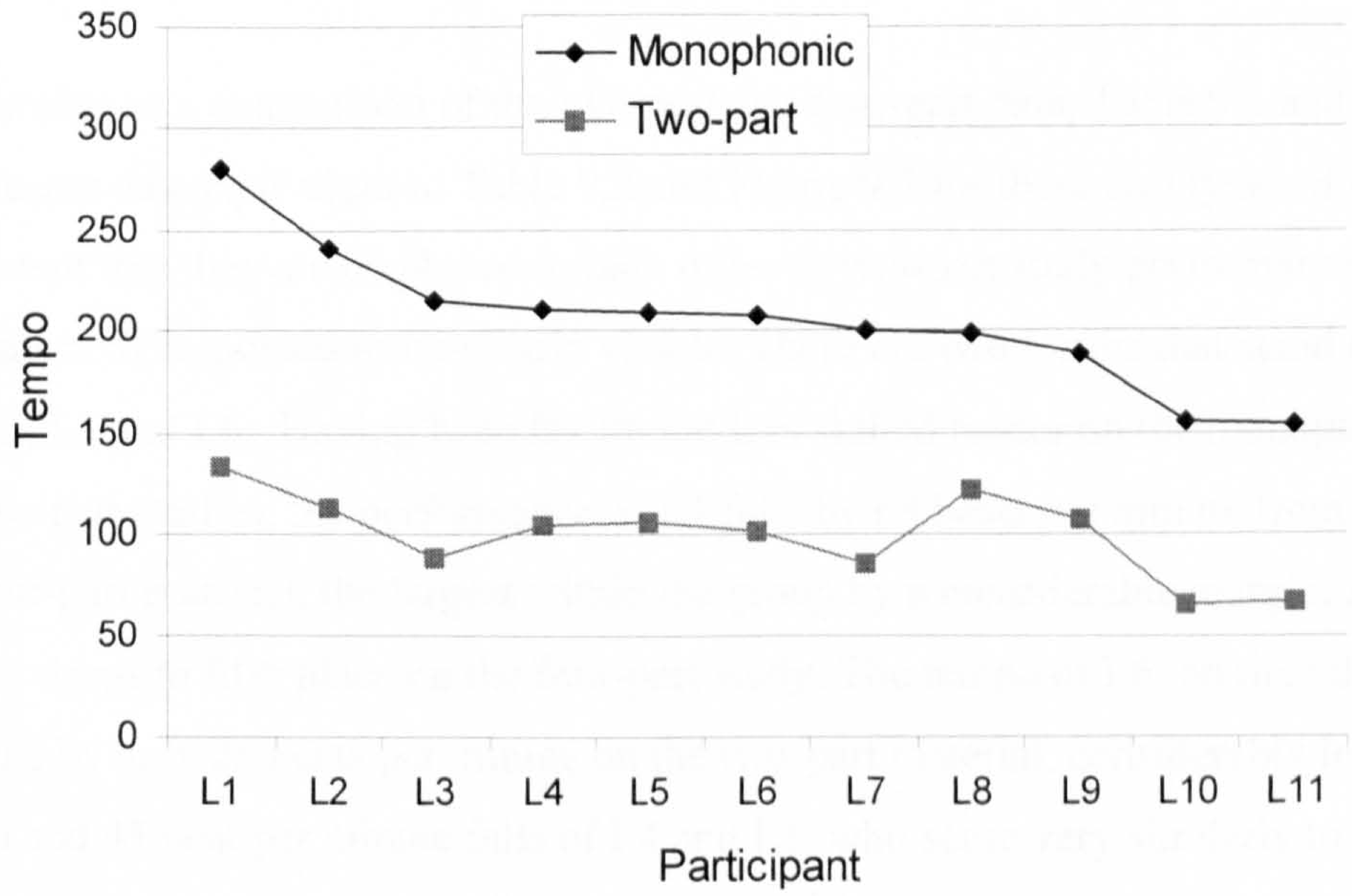


Figure 9.2

Mean asymptotic tempo (beats per minute) for individual less skilled participants performing monophonic and two-part notation, in descending order of monophonic tempo value (combined across structure and key) the two-part study together with the difference between individual pairs of scores. Three of the two-part results particularly stand out as not fitting the pattern of the monophonic distribution: L3, L8 and L9. L3 scores similarly to L4, L5 and L6 on the monophonic study (their results are 215, 211, 208 and 208 beats per minute respectively), but L3 demonstrates a considerably greater fall in tempo than they do with two-part material - a decrease of 126 beats per minute compared to their 103 and 106 beats per minute. Indeed, L3 drops from being the third fastest less skilled reader on the monophonic study to being the fourth slowest on the two-part study. The performance of L8 and L9, in contrast, improves relative to other readers with two-part material. Their monophonic performance results are similar to L6 and L7 (L6 = 208, L7 = 200, L8 = 199, L9 = 188 beats per minute), but the decline in their tempo with two-part material is considerably less (difference between monophonic and two-part results: L6 = 106, L7 = 114, L8 = 77, L9 = 81 beats per minute). In consequence, L8 and L9 are the second and fourth fastest in the less skilled group with two-part material. When the results of L3, L8 and L9 are removed, the correlation between monophonic and two-part performance within the less skilled group changes from being non-significant ($r_s = 0.55$, $p < 0.08$), to being highly significant ($r_s = 0.97$, $p < 0.0001$). This evidence is consistent with these three participants being the principal source of performance variation between the two distributions.

Turning to a comparison of the two-part and four-part data, Table 9.3 and Figure 9.3 present data equivalent to Table 9.2 and Figure 9.2 for these two types of material, except that they are displayed in rank order of two-part study performance to make the pattern of responses more clearly visible. There are two results that stand out in Figure 9.3: L1 and L6. Having been the fastest less skilled reader on the monophonic and two-part studies, the performance of L1 falls by 68 beats per minute from two-part to four-part material, the largest within the group by a considerable margin. As a result, she drops to fifth place on the four-part study. The tempo of L6, on the other hand, falls by only 25 beats per minute on the two-part material, considerably less than the 36 and 41 beat per minute falls of L4 and L5, who score very similarly to L6 on the two-part study (L4 = 105, L5 = 106, L6 = 102 beats per minute). L3, L8 and L9 do not repeat their apparently idiosyncratic performance variation between monophonic and

Table 9.3

Mean asymptotic tempo (beats per minute) for individual less skilled participants performing two-part and four-part notation (combined across structure and key levels, rank order of two-part and four-part performance in parentheses)

	Two-part (rank)	Four-part (rank)	Difference
L1	134 (1)	66 (5)	68
L8	122 (2)	90 (1)	32
L2	114 (3)	77 (2)	36
L9	108 (4)	64 (7)	44
L5	106 (5)	65 (6)	41
L4	105 (6)	69 (4)	36
L6	102 (7)	77 (3)	25
L3	89 (8)	50 (8)	39
L7	86 (9)	44 (9)	42
L11	67 (10)	36 (11)	30
L10	66 (11)	41 (10)	24

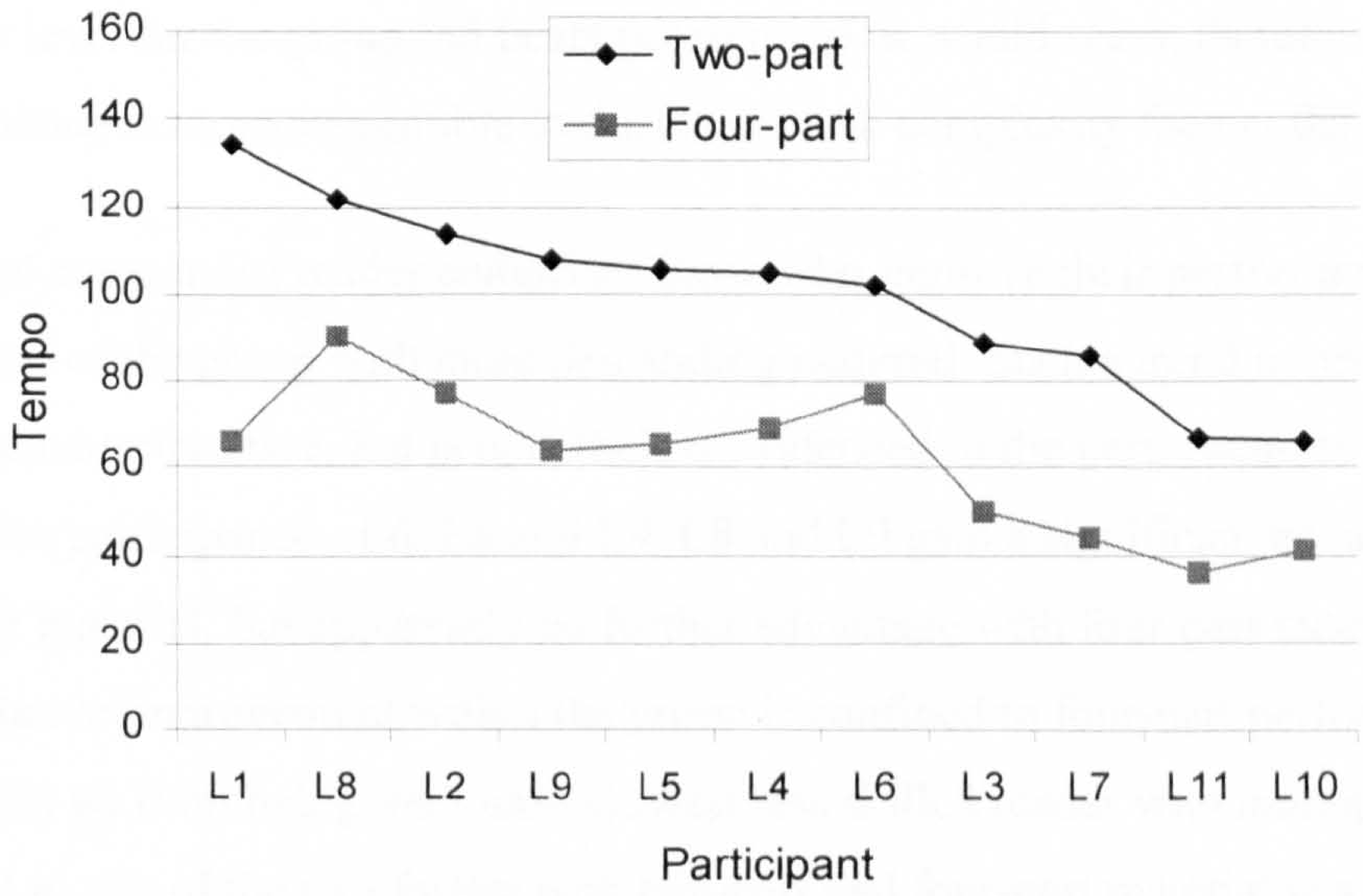


Figure 9.3

Mean asymptotic tempo (beats per minute) for individual less skilled participants performing two-part and four-part notation, in descending order of two-part tempo value (combined across structure and key levels)

two-part material, showing unusually average declines in output from two-part to four-part material (L3 = 39, L8 = 32, L9 = 44 beats per minute). This average decline, however, in combination with L1's unusually large one, is sufficient to make L8 the fastest less skilled reader on the four-part study. When the results of L1 and L6 are removed, the correlation between two-part and four-part performance within the less skilled group increases considerably, from $r_s = 0.76$ ($p = 0.007$) to $r_s = 0.95$ ($p < 0.0001$).

These results indicate three general patterns of behaviour amongst the less skilled group. First, there are the majority of readers (L2, L4, L5, L7, L10 and L11), who perform similarly, in relative terms, on all three studies. Second, there are readers who show a particularly large decline in performance with more complex materials compared to the group overall (L3 and L1). The fact that L3's large decline is only between monophonic and two-part performance suggests that her difficulties may lie with dual-stave performance factors non-specific to the complexity of an individual task, for example a general problem with coordinating hands together motor output. L1 performs within the skilled range with monophonic notation, and although she retains first position amongst less skilled participants on the two-part study, the fact that she now falls outside of the skilled range here would seem to indicate a relative decline in her ability. On the four-part study she drops to fifth place, scoring at only an average level for the group (66 beats per minute). It would seem, therefore, that her performance is more susceptible to dual stave task complexity factors than that of L3.

The final category of reader comprises those who *improve* their performance relative to the rest of the group with more demanding material. Such a trend in results does appear counterintuitive, but is nonetheless evidenced in the performances of no less than three participants – L6, L8 and L9. L8 and L9 gain a significant advantage with two-part material, but apparently no further advantage with four-part material, while L6's relative improvement within the group is confined to four-part performance. That L8 should go from being the fourth slowest less skilled reader with monophonic material to one of the two fastest with two-part and four-part material is a particularly surprising and indeed puzzling result. One explanation, though, could be that this participant had an output related problem preventing her from achieving the tempi of faster less skilled readers on the monophonic study, something that would not have been limiting in the slower context of two-part and four-part performance. A similar explanation might also be proposed for the relative performance gain of L6 with four-part material. However, a possible alternative account is suggested by the fact that this participant (i.e. L6) was a church organist, whose considerable experience with playing hymn-tunes may have provided him with a performance advantage at this analogous experimental task. On the other hand, L8 was also a church organist and there appears to be nothing notably superior about her four-part performance compared with her two-part performance.

9.2.3 Skilled participants

As with the less skilled group, variation between monophonic and two-part performance will be considered to begin with, Table 9.4 and Figure 9.4 displaying the mean tempo data of individual participants for these two types of material at maximum effective preview. From Figure 9.4 it would appear that there are three two-part results in particular that stray from the general trend of the monophonic data: S3, S8 and S10. S3 performs similarly to L3 of the less skilled group, her performance declining from the third fastest on the monophonic study to third slowest on the two-part study. The difference between her monophonic and two-part performance is 171 beats per minute, compared to differences of 114 and 104 beats per minute for S2 and S4, who performed very similarly to her on the monophonic study (S2 = 338, S3 = 333 and S4 = 321 beats per minute). A less dramatic decline in performance is seen with S8, who performs 121 beats per minute more slowly on the two-part study, compared to the 90 and 97 beat per minute fall of S6 and S7, who both score within 4 beats per minute of her on the monophonic study (S6 = 272, S7 = 270 and S8 = 268

Table 9.4

Mean asymptotic tempo (beats per minute) for individual skilled participants performing monophonic and two-part notation, in descending order of monophonic tempo value (combined across structure and key, rank order of two-part performance in parentheses)

	Monophonic	Two-part (rank)	Difference
S1	374	233 (1)	142
S2	338	223 (2)	114
S3	333	162 (8)	171
S4	321	218 (3)	104
S5	309	182 (4)	127
S6	272	175 (7)	97
S7	270	180 (5)	90
S8	268	148 (10)	121
S9	255	162 (9)	93
S10	247	179 (6)	69

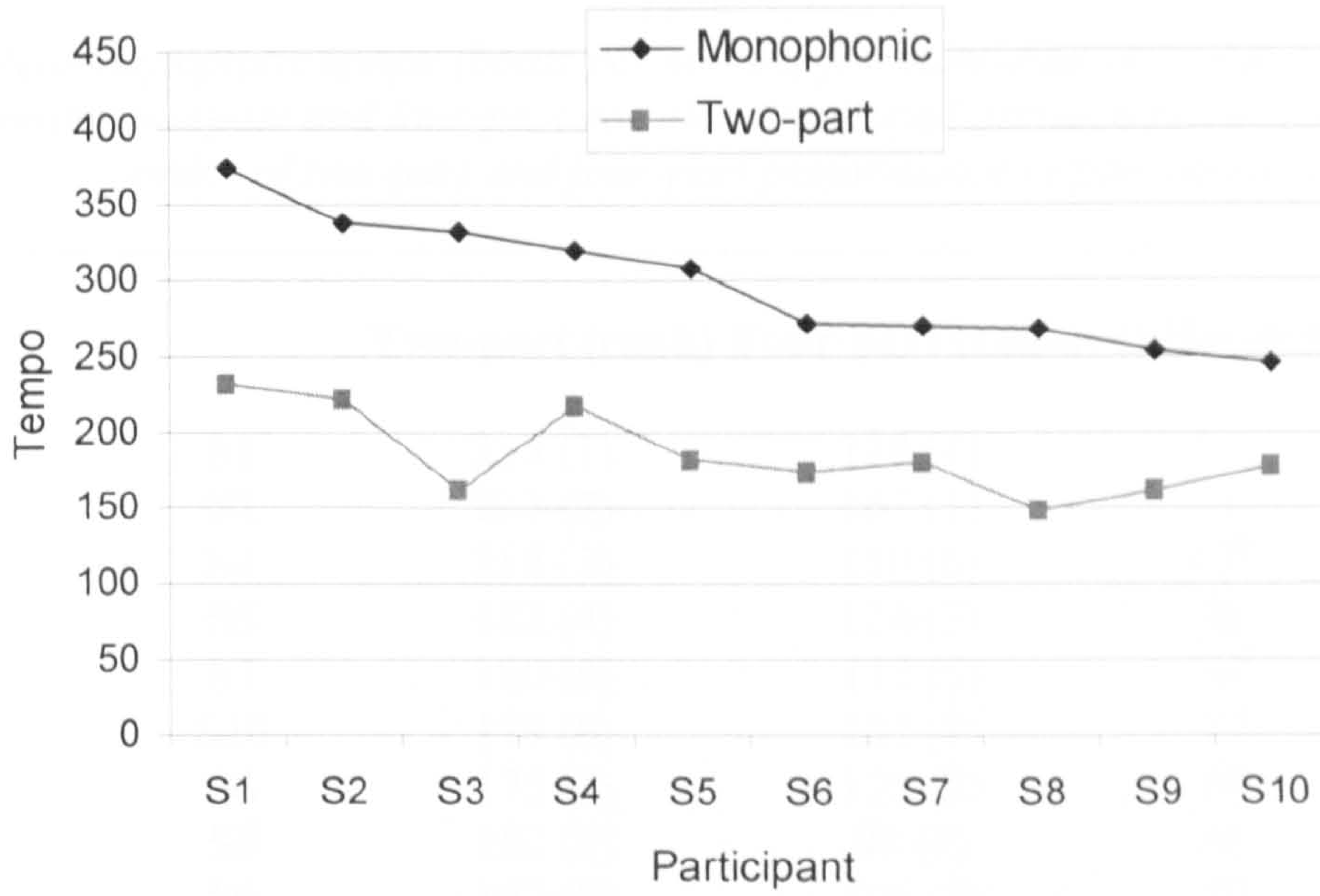


Figure 9.4
Mean asymptotic tempo (beats per minute) for individual skilled participants performing monophonic and two-part notation, in descending order of monophonic tempo value (combined across structure and key)

beats per minute). This makes S8 the slowest performing skilled reader with two-part material. Finally, the difference between S10's performance on the monophonic study and the two-part study is, at 69 beats per minute, the smallest of the entire skilled group. In consequence, she moves from being the slowest of the skilled group on the monophonic study to sixth fastest on the two-part study; in fact, her tempo of 179 beats per minute here is only 3 beats per minute slower than the fourth fastest, S5. When the results of S3, S8 and S10 are removed, the correlation between monophonic and two-part performance within the skilled group increases considerably, from $r_s = 0.65$ ($p = 0.04$) to $r_s = 0.94$ ($p = 0.002$).

Turning finally to the skilled two-part and four-part data, Table 9.5 and Figure 9.5 present the individual participant data at maximum effective preview for these materials. The data are displayed in rank order of two-part study performance to make the pattern of responses more clearly visible. A particular source of variation between the two distributions appears to be the data for S1 and S4. The performance of

Table 9.5
Mean asymptotic tempo (beats per minute) for individual skilled participants performing two-part and four-part notation (combined across structure and key, rank order of two-part and four-part performance in parentheses)

	Two-part (rank)	Four-part (rank)	Difference
S1	233 (1)	118 (4)	115
S2	223 (2)	167 (1)	56
S4	218 (3)	110 (6)	107
S5	182 (4)	124 (3)	58
S7	180 (5)	112 (5)	68
S10	179 (6)	102 (8)	77
S6	175 (7)	125 (2)	50
S3	162 (8)	98 (9)	63
S9	162 (9)	104 (7)	57
S8	148 (10)	93 (10)	54

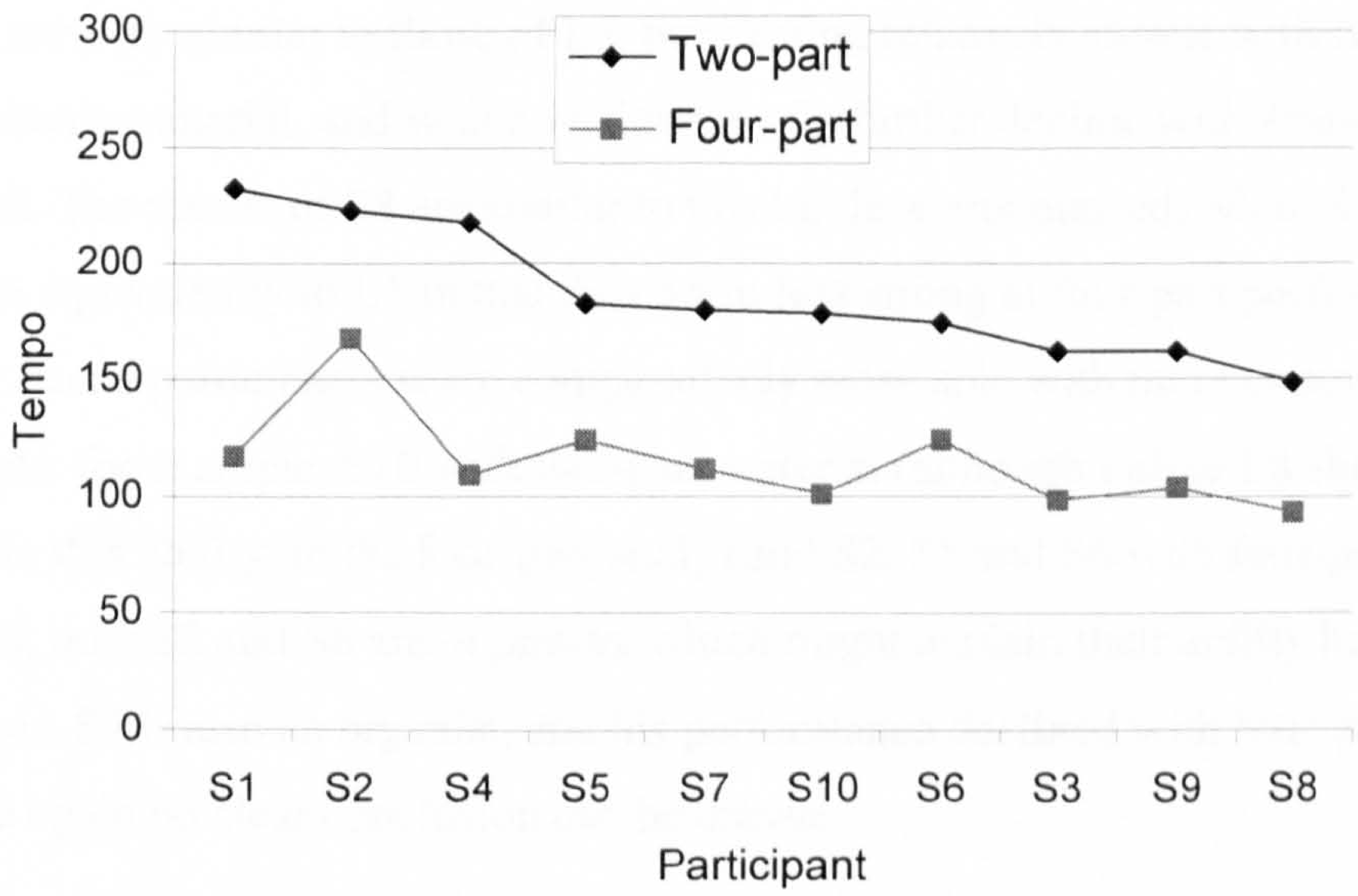


Figure 9.5
Mean asymptotic tempo (beats per minute) for individual skilled participants performing two-part and four-part notation, in descending order of two-part tempo value (combined across structure and key)

these skilled readers declines to rather average levels on the four-part study (118 and 110 beats per minute respectively), the former dropping from first place on the two-part study to fourth place on the four-part study, and the latter from third to sixth place. Meanwhile two of the central, average performing group of four participants on the two-part study (S5, S6, S7 and S10 with scores between 175 and 182 beats per minute) achieved faster four-part performances than both S1 and S4 (S6 scoring 125 beats per minute and S5 scoring 124 beats per minute) making them the second and third fastest performers respectively on the four-part study. Another of this group, S10, shows less facility with four-part reading, attaining only 102 beats per minute, which is the group's third slowest score. S2, like S5 and S6, demonstrates a particular skill with four-part material, but his tempo of 167 beats per minute would seem to be in an entirely different league from theirs: his performance is 42 beats per minute (40%) faster than that of S6, the second fastest skilled group member.

The skilled group data provides evidence of the same three categories of reader found earlier amongst the less skilled group. First, there are those like S9 and S7 who seem broadly equivalent in their ability across the different performance materials, and show little variation in rank order. Second, there are readers whose performance declines relative to the wider group with more complex notation. For example, S3's results are very similar to those of L3, performing relatively slower with two-part than monophonic material, and with no evidence of a further decline with four-part material. The results of S8 are similar to this but less pronounced, while S1 and S4 perform equivalently to L1 in that they seem less strong at four-part performance. Finally, some participants seem comparatively *more* able with more complex materials: for example, S10 with two-part material (although unlike L8 she does not maintain this ability on the four-part study) and S2, S5 and S6 with four-part material. Like L6, both S2 and S6 are organists, which might explain their ability here. However, S4 is also an organist, and his performance declined with four-part material, so once again no clear conclusion can be drawn.

To really understand the relevance of these data to the real-world activity of sight-reading, clearly more research into this required. For example, with all members of the skilled group considering themselves confident sight-readers, are faster performances within the group irrelevant to actual sight-reading requirements? Or are they linked to more skilled levels of performance? However, some limited insight into these issues can perhaps be provided by the current research. Amongst the skilled group, only four participants, S1, S2, S3 and S4 make their living primarily from their sight-reading ability. S1, S2 and S3 are professional accompanists (S2 is also a professional organist) and S4 is a resident cathedral organist. The fact that these individuals are the fastest performers of monophonic material may indicate that this task is testing something that is integral to fully professional levels of sight-reading performance. It is not clear what this might be, though - it would seem unlikely that the important factor is the achievement of speed itself, because the speeds attained - well over 300 beats per minute - would seem beyond typical sight-reading requirements. Whilst S1, S2 and S4 continue, by a considerable margin, to be the best performers with two-part material, the performance of S3 declines here to one of the slowest speeds within the group (eighth place). On the four-part study, only S2 and S4 remain among the top four scoring readers (first and fourth place respectively), S3 being in ninth place, and S4 in sixth place.

These results indicate that average or even poor performance within the context of the skilled group on these two latter tasks is not necessarily a hindrance in professional terms. Perhaps this is, as mentioned earlier, simply evidence of performance beyond a basic skilled level on the two-part and four-part studies being superfluous to actual sight-reading requirements. However, it is also possible that for some musicians, skill at these tasks may be more fundamental to their sight-reading ability than others. During informal conversation with S1, he commented that much of his own work involved what he termed 'bluffing': the use of musical and stylistic knowledge to create a partially improvised rather than completely accurate performance of the notation. S2 on the other hand claimed not to resort to this, priding himself on the accuracy of his performances. Perhaps the greater consistency of S2's performances across the different experimental materials is reflective of this – the highest level of performance on each being necessary to support his particular approach to sight-reading. Obviously, this proposition is only a speculative one, but it raises an important issue for future research to consider. It must also be borne in mind that many other factors may be important in determining overall sight-reading skill and so deserve research attention, for example, the ability to transcribe rhythms and read confidently in more complex keys.

9.2.4 Summary of analysis and conclusion

Considering differences between the skill groups to begin with, the evidence indicates that for two-part and four-part materials the results of participants form two discrete groups reflective of skill group membership. This is typically the case for monophonic material as well, except for two less skilled readers who perform within the skilled reader range. Generally, therefore, performance on a particular category of material can be considered predictive of skill group. However, both skill groups exhibit substantial within-group variation. Although the source of most of the rank variation between each pair of skill group distributions analysed typically lies with only two or three individuals, overall only two skilled and six less skilled participants display a consistent and strong relationship both between their performances with monophonic and two-part material, and between their performances with two-part and four-part material. All the other 13 participants show signs of some important variation in their responses, typically, but not always, limited to one particular type of material. Further work is clearly necessary to confirm the validity of these findings, but the evidence suggests that that sight-reading is a more complex activity than has generally been considered by previous research, requiring more sophisticated explanation. The findings also provide support for the idea proposed in the literature review that the measurement of sight-reading ability should be multi-dimensional in nature i.e. take into consideration the possibility of variation in skill across different component tasks by measuring them separately.

This analysis has indicated that the undertaking of more specific investigation into individual readers' abilities and difficulties, begun here, would seem a fruitful and indeed necessary direction for sight-reading research to continue along; a more detailed knowledge of the 'anatomy' of the skill is surely fundamental to the development of a more comprehensive theoretical understanding. The current analysis has not been able to provide any further progress in actually explaining the skill differences that have been observed. This research will now move on to the investigation of perceptual and motor sub-skills to see if they are capable of providing further insight into these matters.

Chapter 10

Perceptual sub-skill study

10 Perceptual sub-skill study

This chapter is divided into three main sections:

- A description of the experimental methodology
- Results and statistical analyses
- Discussion of research findings

10.1 Methodology

10.1.1 Development of experimental design and technology

The rationale behind this work has been briefly discussed in Chapter 3, but will now be considered in more detail. In the context of the study, the term 'perception' is viewed in its broadest sense, encompassing basic visual perception of the score, together with transcription and cognitive processing as far as the interface with motor output mechanisms. Recent understanding, dominated by the patterning account, has typically considered that the source of the difference in sight-reading ability of skilled musicians lies principally within the realm of perceptual processing, as just defined. However, such an understanding has been largely based upon experimental tasks involving a significant motor component (for example, Sloboda's eye-performance span studies (Sloboda, 1974, 1977)), with researchers and reviewers apparently failing to consider that the motor component itself might be a significant source of variation in the perceptual measures recorded. A priority for research is therefore to gain a more narrowly focussed insight into the extent to which skilled and less skilled readers vary in just their perceptual abilities. To achieve this, there is a need for work that isolates perceptual activity more from motor influence, as the study discussed in this chapter has attempted to do. Exploring perception using the same participants as in the complete task preview studies, together with gaining further understanding of their sight-reading related motor abilities (see chapter 11), also provides the opportunity for insights to be gained into the relative importance of the two sub-skill areas in determining participants' abilities at the complete sight-reading task.

Perceptual activity obviously cannot be completely isolated experimentally from motor output; some form of motor involvement is necessary so that the completion of an experimental task can be signified and the accuracy of perception gauged. One potential problem of stripping perception of its normal motor context is that it may may lead to fundamental changes in the processes involved, meaning that recorded measures are limited in their relevance to the performance of the full task. A particular case in point is one of the studies carried out by Waters *et al.* (1998a) in which they attempted to measure the speed of perception and transcription of individual notes by requiring participants to name them out loud. It would seem questionable that such conscious naming of notes is fundamental to perceptual processes within sight-reading performance. A complete resolution of this problem is clearly unattainable, since it is unrealistic to expect that an experimental procedure could be devised that would directly replicate and measure the actual mechanisms involved in sight-reading prior to output. However, if a procedure were to involve a constraining of perceptual activity to sub-skills that are at least essentially relevant to the sight-reading process, a degree of validity could be achieved. What is required is a task that involves the meaningful processing of musical notation, but demands only a limited motor output that is perceptually neutral i.e. that does not itself impose any particular direction on processing.

Waters *et al.'s* pattern matching studies (1997, 1998a), discussed previously in the literature review, would appear to be more successful than their note-naming experiment in terms of the neutrality of the motor activity involved. These studies required participants to compare two sequences presented as standard notation, and to press one of two buttons depending upon whether the two sequences were identical or varied by a single note. This use of error detection encourages meaningful perceptual processing of the experimental material, something that is open to evaluation through a consideration of error rates. A correct matching of two identical sequences can generally be considered to have involved the explicit processing of the entire sequence, and the time taken for this a relative measure of the speed of the perceptual sub-skills involved in the task. In these experiments, therefore, the pushing of the button is clearly not the goal of the perceptual activity, but simply a means of indicating its speed and accuracy. There are a couple of other advantages of the pattern-matching approach. Firstly, because trials involve groups of notes, the time taken for the planning and execution of the motor component is only small in relation to overall response time. Secondly, the motor component does not interfere with perception-related processes. For example, experiments to explore the quantity of stimulus material that musicians can perceive have sometimes required them to write down their responses (Waters *et al.*, 1998). Not only is there a greater likelihood of natural memory decay because of the longer time required for such responses, the extended level of attention required might possibly serve to accelerate it.

A problem with Waters et *al's* empirical implementation of the pattern-matching approach is that their perceptual task was not as tuned to the realities of sight-reading activity as it might have been. Transcription was not an essential component of the comparison procedure, and so one cannot be sure of the extent to which the strategies and representations employed by participants were musically relevant. One way of ensuring this would be to require participants instead to compare a sequence of standard notation with its transcribed keyboard representation, presented visually. Although such an adaptation would require a response based upon musical knowledge, it still has the weakness that the strategies employed by participants may not relate to their normal sight-reading related perceptual processing. For example, there is nothing to stop them using a completely letter-name based technique. Such a modification to Waters *et al.'s* method is therefore clearly not ideal, but it would at least represent a step forward in the research process.

In relation to the current thesis, therefore, it was felt that Waters et *al.'s* pattern-matching methodology, incorporating the above modification, provided a worthwhile means of investigating the perceptual ability of musicians. Responses could be measured in the context of different sequence lengths and types of musical structure, enabling detailed exploration of the skill patterns of the two participant groups that could also be directly related back to results from similar conditions within the controlled preview studies. The exploration of these conditions also allows for the possible testing of the proposals of the patterning account within a specifically perception related context.

Despite the apparent advantages of the pattern-matching approach, there are limitations that need to be borne in mind. Although the method minimises variation in the data from hand and arm motor activity, significant variation of non-perceptual origin might nevertheless result from the eye-movements involved in the comparison process. There is no easy resolution to this, but there is perhaps a case for arguing that the issue is not particularly important considering how integral eye-movements are to visual perceptual activity. The approach can also only provide a general measure of overall perceptual sub-skill. Although some insight into sensitivity to musical structure may be achievable, the relative importance of the other perception-related skill components in constraining responses cannot be determined. Such components include, for example, skill at generic visual perception, note encoding, transcription, and the mental comparison of stimuli.

To undertake the research, an experimental tool was required that would present on a computer screen a stave containing a sequence of notation together with a graphical representation of a section of piano keyboard upon which the keys corresponding to the notation could be highlighted. Also required was a timed means of responding to sequences as either matching or non-matching. As was the case with the controlled preview experiments, no commercially available software was available that could be adapted to carry out the particular tasks required, and so a new computer program was designed and written in the C language.

10.1.2 Materials

10.1.2.1 Basic design overview

No formal pilot study was carried out. The process of software development required regular testing of the technology, and this was undertaken using musicians of a range of ability. From this process, the final methodology was honed in relation to the screen design, choice of materials and general running of the experimental procedure.

In order that a more complete understanding of foundational perceptual issues might be achieved, it was decided to focus on a more in depth study of monophonic materials, rather than a superficial examination of all of the notational complexities used in the controlled preview research. Because a principal purpose of this experiment was to inform the results of the monophonic preview study, the experimental materials were designed to reflect as closely as possible the content of that study's sequences. Sequences were therefore written using the same range of the treble clef stave - middle C to the G at an interval of a 12^{th} higher - consisting only of crotchets. Also, except for the exclusion of a condition of key, they involved the use of equivalent independent variables. A key condition was not appropriate considering that sequences were too short to enable any sense of a home key to be defined.

Only continuously ascending note sequences were used. It quickly became apparent during the informal piloting that participants found material that rose and fell, or simply fell, confusing to perform, something which led to significant delays in the comparison procedure, and which it was considered compromised the data as valid measures of basic perceptual ability. In the case of sequences that rose and fell, the notes appeared in a different order on the keyboard than on the stave, meaning that comparisons did not necessarily involve groups of adjacent notes, as had been the case in Waters *et al.'s* original design. Figure 10.1 provides a screen shot demonstrating this issue. Although with continuously descending sequences the comparisons could take place with adjacent notes, the problem was that a group of notes on the stave was always in reverse order to the corresponding group on the keyboard graphic, making the matching process far from intuitive. An example of this is shown in Figure 10.2. Figure 10.5 provides an example of a non-matching rising sequence, for comparison. Although the use of only ascending sequences imposed some limits on the variety of material that could be tested, it would appear an acceptable compromise.

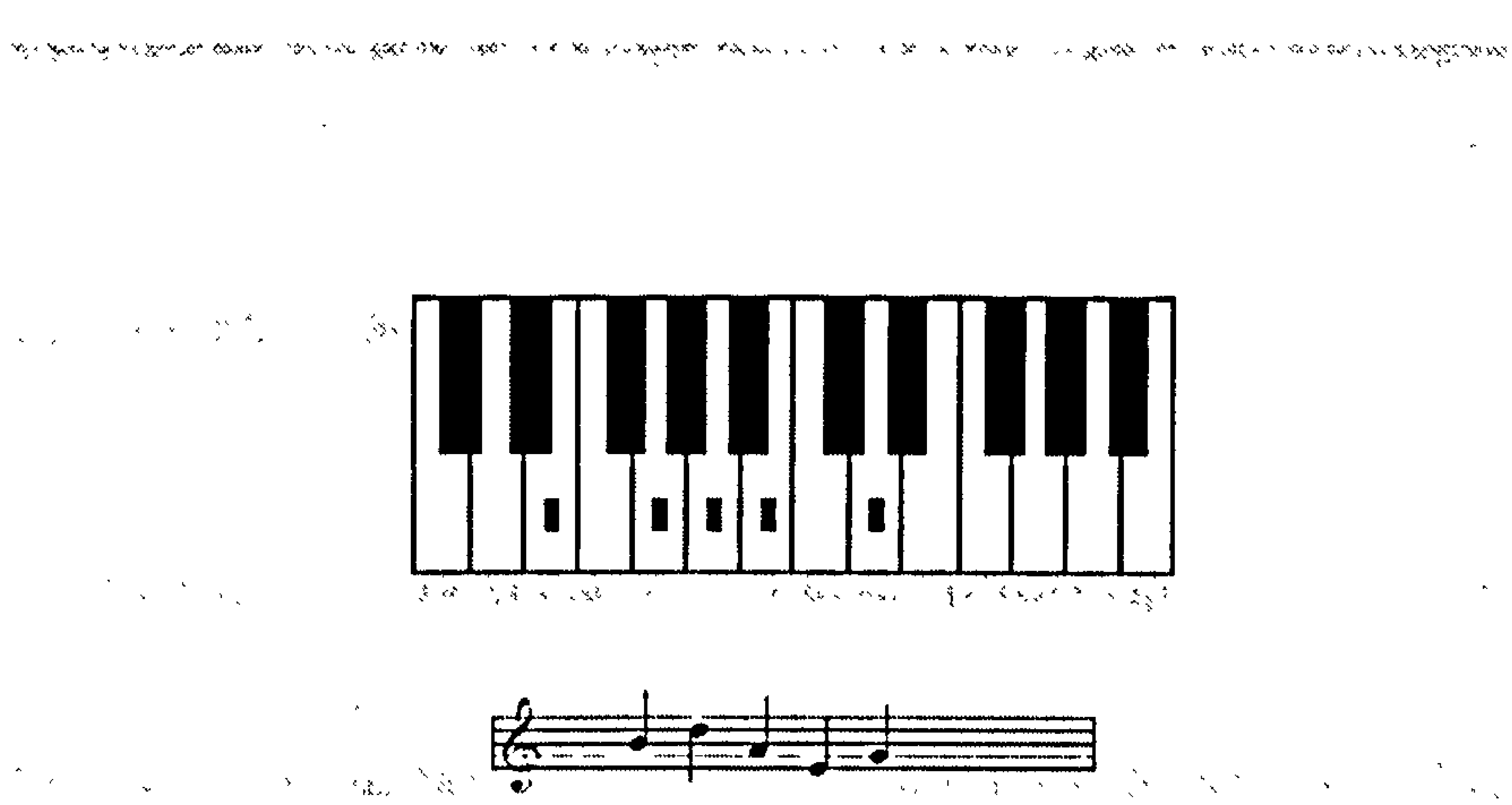


Figure 10.1
Screenshot showing a rising and falling sequence

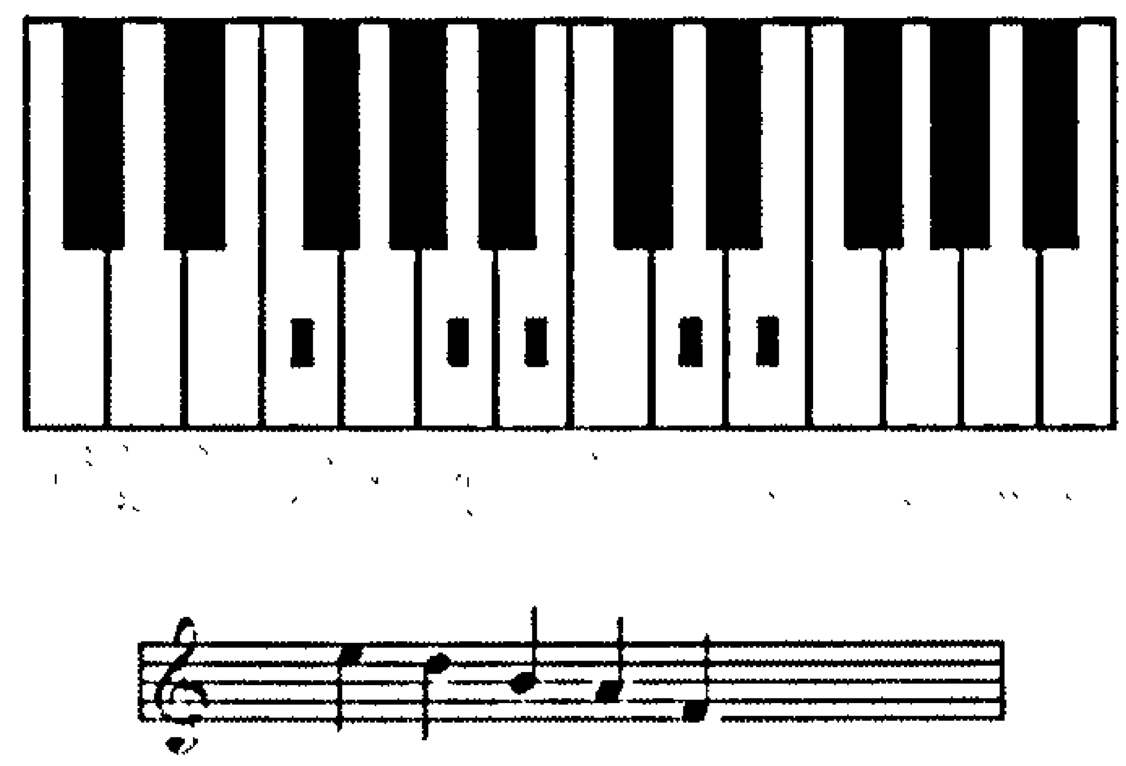


Figure 10.2
Screenshot showing a falling sequence

10.1.2.2 Factors to be studied

Sequences were designed to explore the influence of two factors on perception: sequence length and musical structure.

Sequence length

Experimental materials comprised 5 different levels of sequence length: 1, 2, 3, 4 and 5 notes. This range reflects the preview sizes that had generally been found to result in significant performance speed variation during the informal piloting of the monophonic controlled preview software, enabling the results of this current study to be directly compared to those of the equivalent preview size levels on the monophonic controlled preview experiment. Participant responses to these sequence lengths may also enable some insight to be gained into the whether skilled readers use larger perceptual units for their processing than less skilled readers.

Structure

To enable this experiment to test the validity of the patterning account and also to inform the results relating to the effect of structure in the monophonic preview study, both structured and unstructured materials were used for sequences of 3, 4 and 5 notes in length. Obviously, the involvement of structure is not relevant for 1-note sequences, and would have been necessarily ambiguous in relation to 2 notes. Given the limited length of sequences, it is also clear that, unlike the controlled preview work, structure was limited to being local in nature. The full range of white notes in the treble clef stave from C to the top G was represented in 1-note sequences, as well as F# and Bb. For 2-note sequences, intervals from a 2nd to an octave were employed. Materials for 3, 4 and 5-note categories were designed to provide a structural distinction as similar as possible to the monophonic preview study; indeed the aim was that they should be like excerpts from these sequences (given the limitation that all sequences were required to be ascending). The basic note collection was the same as for the preview study – all the white keys, F# and Bb. It was also necessary to make some use of C# and Eb to provide unstructured material with more variety of content. For example, tritones were useful in the design of these materials, but with only a limited number available in the keys of C, F and G there was the danger of their being over utilised. The extra black notes were not used chromatically: all material was diatonic to be consistent with the monophonic preview study sequences. Structured and unstructured material at 3, 4 and 5-note sequence lengths were written according to the following specifications.

Structured material

Structured materials contained the foundational elements of tonal melodic structure used in the monophonic preview materials and appropriate to rising melodies: close use of major and minor broken triads and passing notes. The sequences employed either no passing notes, or just one, and to provide the greatest structural integrity, each sequence implied a single triad harmonically. 3 and 4-note sequences reflected the possible permutations of broken triad inversion and passing note use. With 3-note material the melodic range was from a 3rd to a 6th, the former representing two adjacent triadic notes with an intervening passing note, the latter the range of an inverted triad in close position. With 4-note sequences the range was from a 5th to an octave, the former representing a root triad in close position together with a passing note, and the latter, four adjacent triadic notes in close position. With 5 notes, all sequences were an octave in range. A larger range was not used here because it was not typical of the monophonic preview study materials. Sequences contained either no black notes or just a single one, to keep the visual complexity of the notation to a minimum. Examples of structured sequences are shown in Figure 10.1, and the complete collection is presented in Appendix 2.

Figure 10.3

Examples of 3, 4 and 5-note structured sequences

Unstructured Material

As with the controlled preview study, structure clearly could not be entirely removed from sequences given the limited note collection that was used. The unstructured material was devised to disrupt, as much as possible, the tonal patterning found in the structured sequences. This was achieved in two ways. Firstly, sequences were written that required more than one triad to achieve a simple, triadic-based harmonisation. Either two or three changes of triad were employed for the 3-note sequences, and at least three for the 4 and 5-notes sequences. Secondly, flowing melodic movement was disrupted by not using passing notes. The intervallic range encompassed for each of the sequence levels was the same as for structured material, except an interval of a 7th was used in the 3-part sequences to help accentuate the sense of disruption. Using equivalent ranges was considered important so that the experiment would be measuring as much as possible the perception of musical structure, not variation due to superficial notational elements. The number of sequences containing black notes was matched, within sequence length, to that of the structured materials. Examples of unstructured sequences are presented in Figure 10.2, and the complete collection can be found in Appendix 2.

Figure 10.4

Examples of 3, 4 and 5-note unstructured sequences

<u>10.1.2.3 Further design issues</u>

130 sequences were used in the experiment - 88 matching stave/keyboard pairs and 42 non-matching pairs. It was felt that having an equal number of matching and non-

matching pairs was unnecessary; all that was needed was for the latter to be sufficiently represented to encourage participants to process the entire contents of matching sequences for their responses, rather than be tempted to guess the outcome. This enabled a saving of experimental time, because the results from non-matching pairs were not useful empirically - there was no way of determining the extent to which non-matching pairs had been processed prior to error detection.

Of the 88 sequences used for matching experimental pairs, there were 14 1-note and 2-note sequences, and 20 each of 3, 4 and 5-note sequences. The reason for the smaller number of 1 and 2-note sequences was that they did not have an associated structure condition, and their shortness meant that there was less variety of content to be explored. For 3, 4 and 5-note sequence lengths, the 20 sequences employed for each consisted of 10 structured and 10 unstructured. Turning to the 42 sequences used for the non-matching stimulus pairs, a similar distribution of sequence length and structure levels was used: 6 sequences each at 1 and 2-note sequence lengths, and 10 at 3, 4 and 5-note sequence lengths (5 structured and 5 unstructured). These non-matching pairs differed by only a single, randomly chosen note. The intervallic value for the difference was either a tone or semi-tone; it was felt that if the difference was made as obscure as possible, it would encourage a more attentive processing of the materials. Since the non-matching sequences are not relevant to the analysis, they have not been included in Appendix 2.

All the 3, 4 and 5-note sequences used were evaluated by the independent adjudicator in relation to their musical structure. His task was simply to confirm whether or not the structured and unstructured sequences complied with the design brief described earlier. In relation to structured sequences, this involved checking that sequences were based on only a single broken triad, and that correct usage of passing notes had been maintained. For unstructured material it involved ensuring, firstly, that sequences necessitated sufficient changes of triad i.e. at least two for 3-note sequences and at least three for 3 and 4-note sequences, and, secondly, that no passing notes had been employed. The adjudicator confirmed the categorisation of all materials.

10.1.3 Apparatus

The experimental tasks were performed on a Pentium 4 laptop computer with a 14" screen, running the Windows Millennium operating system. The computer was situated on an adjustable height table, with participants sitting at a distance from the screen that made both viewing and the use of the attached keyboard as comfortable as possible. The researcher ran the experiment from behind the participant, out of their view, using an external cordless keyboard.

The screen interface of the software is shown in Figure 10.5. It consists of a short, black treble clef stave displaying standard music notation, which is situated in the centre of the screen on the horizontal axis, slightly below the vertical mid-point. No time signature or key signature is displayed, and the size of the notation is equivalent to that used in the monophonic controlled preview study. Directly above the stave and slightly above the vertical midpoint, is a graphical representation of a two-octave section of piano keyboard, with middle C as its lowest note. Each of the keys can be highlighted with a white-framed, red oblong marker to indicate the particular keyboard sequence requiring comparison with the stave notation display. Red was

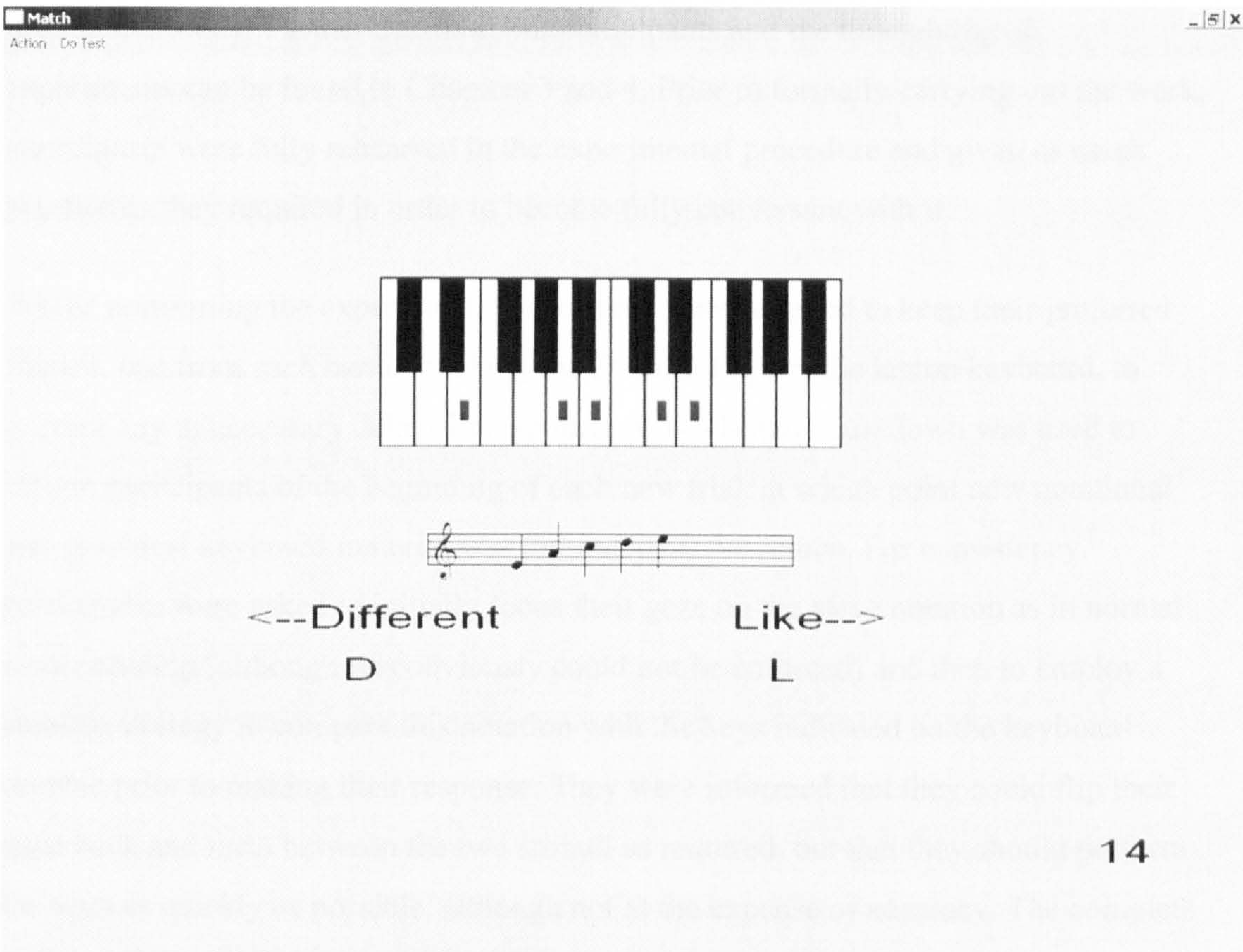


Figure 10.5
Screen interface of the pattern matching software
(showing a non-matching pair of 5-note sequences)

chosen because it stands out clearly against both black and white keys. The screen background is pale grey to make for a comfortable, glare-free visual environment. Each individual test consists of either a matching or non-matching combination of notes and keys presented simultaneously on the screen for participants to respond to. If the notes and keys are considered a perfect match, the L ('Like') key on the computer keyboard is pressed with the right hand. In the case of a mismatch, the D ('Different') key is pressed instead, with the left. (This information is also presented graphically on the computer screen as a reminder to participants). These two keys sit conveniently and comfortably under the forefingers or middle fingers of the hands. The computer then records which key has been pressed together with the time (in milliseconds) that has elapsed since the beginning of that particular display of material.

10.1.4 Procedure

The same participants who had performed the controlled preview experiments were used in this study. Further details about participants and the timetabling of experiments can be found in Chapters 3 and 4. Prior to formally carrying out the work, participants were fully rehearsed in the experimental procedure and given as much practice as they required in order to become fully conversant with it.

Whilst performing the experiment, participants were required to keep their preferred fingers, one from each hand, touching the D and L keys of the laptop keyboard, to prevent any unnecessary delays in response. A 3-2-1-start countdown was used to inform participants of the beginning of each new trial, at which point new notational and graphical keyboard material was presented on the screen. For consistency, participants were asked to initially focus their gaze on the stave notation as in normal music reading (although this obviously could not be enforced) and then to employ a suitable strategy to compare this notation with the keys indicated on the keyboard graphic prior to making their response. They were informed that they could flip their gaze back and forth between the two stimuli as required, but that they should perform the tests as quickly as possible, although not at the expense of accuracy. The complete regime of tests was presented in a different random order for each participant. Between each test, there was a pause of a number of seconds whilst the participant's response was saved to disk. Because of the danger that the large barrage of tests might lead to a loss of concentration, slightly longer rests were given after test numbers 43 and 87 had been performed. Participants were also permitted to initiate a temporary halt to proceedings if required. During the piloting of the software it was found that although participants quickly became accustomed to pressing the key appropriate to their response, they occasionally realised that they had pressed the wrong one by mistake. It was felt important to allow some means of redress if this occurred and therefore participants were allowed to alter their initial response by informing me directly after such an error had been made.

Including time taken for explanation and training, the entire experiment took approximately 30 minutes to carry out. Afterwards, participants were engaged in a short, informal discussion about the representational strategies that they considered they had used, for example, kinaesthetic, auditory, visual, intervallic and alphabetic (note letter names). With participants having already committed considerable time and effort to the study, this discussion could not be a detailed one. Whilst such introspections are not necessarily dependable, it was felt useful to gain at least some small idea of the range of strategies that participants felt they were using in order to assess how these might relate to their skill group membership. The results of this discussion are presented in Chapter 12, together with those of a similar discussion undertaken after the motor sub-skill study (see Chapter 11).

10.2 Results

10.2.1 Introduction

Data relating to three dependent variables have been recorded:

- Response time - the principal dependent variable. This is the time between presentation of stimuli on the computer screen and the pressing of the appropriate key. To provide a more intuitive comparison of results for each level of sequence size, the response time data have been standardised to provide a mean response time per note, dividing the total time taken by the number of notes in the sequence.
- Key press errors - where participants were aware that had mistakenly pressed the wrong response key.

- Matching errors - in which matching pairs were mistakenly considered by participants to be non-matching (which of course may include key press errors that they were not aware of making).

The independent variables are:

- Skill (2 levels: 10 skilled and 11 less-skilled participants)
- Sequence length (5 levels: 1, 2, 3, 4 and 5 notes)
- Structure - relevant to sequence length levels of 3, 4 and 5 notes only (2 levels: structured and unstructured)

Because of the lack of any structure condition at 1 and 2-note sequence lengths, the data analysis for this experiment must be performed using two separate ANOVAs:

1. A one-factor between subjects (skill), one-factor within-subjects (sequence length) repeated measures ANOVA.
2. A one-factor between subjects (skill), two-factor within-subjects (sequence length and structure) repeated measures ANOVA. This ANOVA will be primarily concerned with the role of structure, with sequence length having already been explored in the former analysis.

No statistical analysis has been carried out into errors in view of the fact that only a small number were made by participants.

10.2.2 Principal focal points of the analysis

It would be useful to briefly restate the main issues that this study has been designed to explore. The principal research aim is to find out whether the skilled participants are faster in their perception-related sub-skill than the less skilled participants, something that may help to explain skill difference at the complete sight-reading task. If this is the case, it needs to be considered whether there is evidence that their performance advantage in the sub-skill area is primarily the result of sensitivity to musical structure and the use of larger chunk sizes, as proposed by the patterning account, or whether a modified version of the patterning account might be found to be more appropriate, in which the source of the skilled readers' superior processing lies primarily in their perception of patterning within chunks of similar size to those used by less skilled participants. Alternatively, it might be the case that skilled readers simply have faster elemental perceptual processing skills, and perhaps larger chunk usage, independent of structural influence. It will be particularly interesting to see whether there is again evidence of the less skilled groups' responses being equivalent to those of the skilled group in terms of sensitivity to structure.

10.2.3 One factor between (skill), one factor within (sequence length) ANOVA

As with the controlled preview experiments, the data from both skill groups are positively skewed. However, considering that the two distributions are similar in shape and reasonably homogeneous in terms of variance (skilled SD = 300ms, less skilled SD =380ms), they are generally acceptable for the purposes of ANOVA. Appropriate adjustments (Greenhouse-Geisser) have been made, where necessary, to counter the lack of sphericity within the data, their use indicated by the alternative method of presenting the statistical evidence introduced in Chapter 5, for example, $F_{7, 133}$ [1, 19] = 60.62, $p < 0.0001$. The numbers in subscript represent the unadjusted degrees of freedom, and those in square brackets the degrees of freedom after multiplication by the epsilon adjustment, which in this case has the value 0.143 (to 3 decimal places). The issue of sphericity within repeated measure ANOVAs has already been discussed in Chapter 5, and I refer the reader there for further information. A standard significance level of 5% has been employed throughout the analysis.

The main effect of skill is highly significant (F (1, 19) = 11.25, $p = 0.003$), with the skilled readers demonstrating faster average response times overall. The mean response time per note for skilled participants across all sequence lengths is 715ms (SD = 300ms) and the equivalent figure for less skilled participants is 887ms (SD = 380ms). The other main effect of sequence length is also highly significant ($F_{4, 76}$ [1, 19] = 226.23, $p < 0.0001$). The relevant data are presented in Table 10.1. They reveal what appears to be a quadratic trend, with response time decreasing towards an asymptote with increasing sequence length. There is no significant interaction between the two factors ($F_{4, 76}$ [1, 19] = 0.39, $p = 0.54$), indicating that the significant effect of skill is consistently maintained across all sequence length levels. One would not normally undertake to analyse this non-significant interaction further, but in this case it would be helpful to present the individual group distributions to

Table 10.1

Mean response time per note (ms) for combined skill groups at levels of sequence length (combined across structure, SD in parentheses)

Sequence length (notes)	1	2	3	4	5
Response Time	1235 (397)	883 (362)	665 (225)	627 (195)	596 (202)

Table 10.2

Mean response time per note (ms) for each skill group at levels of sequence length (combined across structure, SD in parentheses)

Sequence length (notes)	1	2	3	4	5
Skilled	1142 (320)	788 (234)	575 (171)	557 (165)	515 (149)
Less skilled	1329 (438)	979 (429)	755 (234)	698 (196)	677 (213)

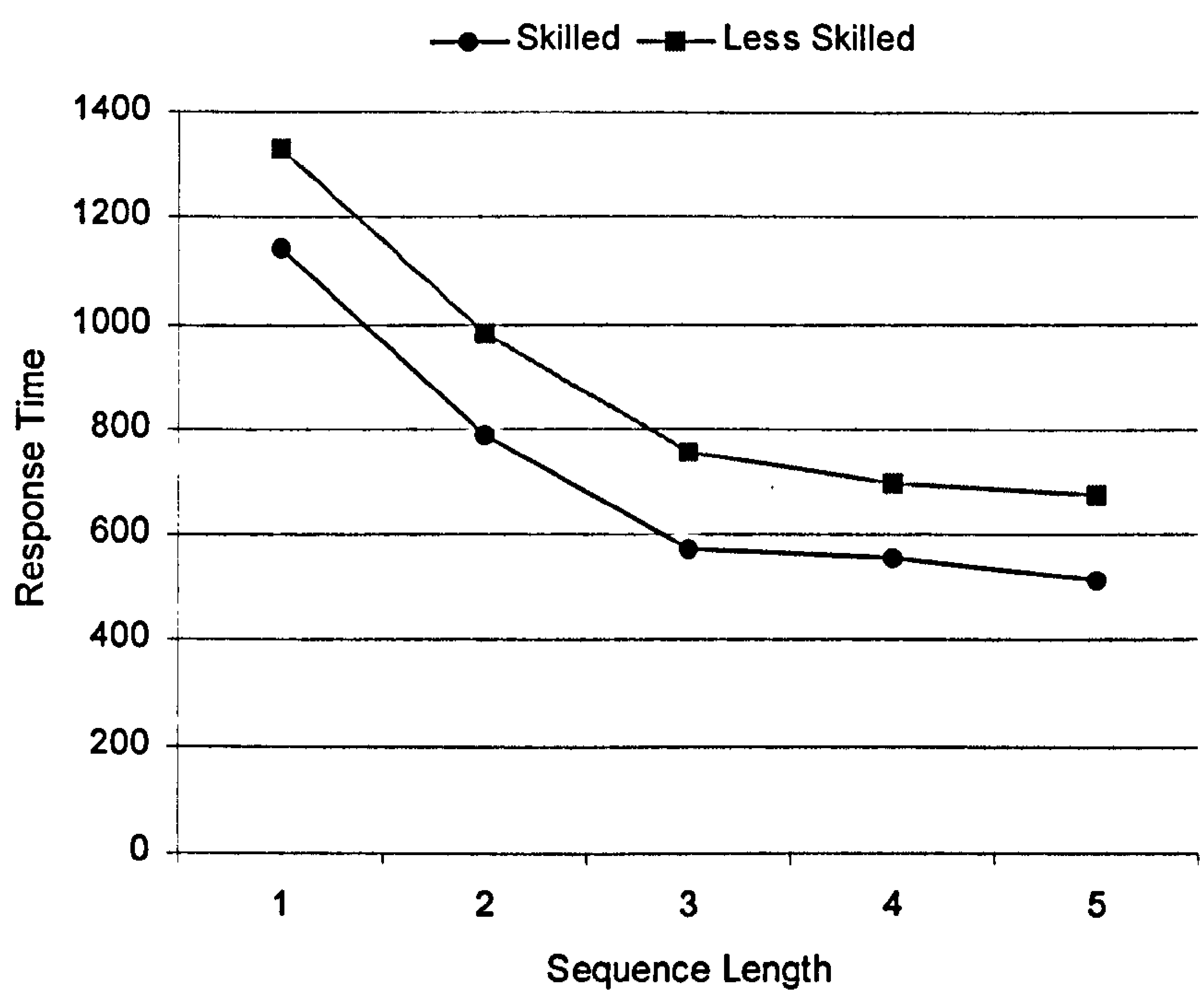


Figure 10.6

Mean response time per note (ms) for each skill group at levels of sequence length (combined across structure) provide a more direct visual impression of the skill effect. The data are shown in Table 10.2 and Figure 10.6.

It is necessary to ascertain whether the effect of sequence length is statistically significant across all levels, or whether there is evidence of an asymptote having been reached with longer sequences. Although there is no significant skill x sequence length interaction, it would seem important to investigate this for each skill group individually, given the central concern of this research to gain insight into how skilled and less skilled readers differ in skill make-up. To accomplish the task, linear contrast analyses have been employed, the same method that was used to determine the preview sizes at which asymptotic performance was reached in the controlled preview studies. As was the case with those studies, this method is considered to be more appropriate than simple statistical comparisons of individual means because it takes into account overall trends in the distributions, spreading the effect of possibly eccentric data. The appropriate weightings for the linear contrast analyses together with the results are presented in Tables 10.3 and 10.4. The results in Table 10.3 indicate that the asymptotic response time for the skilled participants is reached at a sequence length of 3 notes, but considering that weighting 1c is approaching significance, it is possible that the asymptote may be reached at 4 notes. From Table 10.4 it would appear that the asymptotic response time of the less skilled group is reached with sequences of 4 notes. However, it must be borne in mind that weighting 2c is only marginally significant. In the absence of any clear-cut statistical result for either group, perhaps a safe conclusion would be that both the skilled and less-skilled readers reach asymptotic levels of response at sequence lengths of 3 to 4 notes.

Table 10.3

Linear contrast analyses to determine the sequence length at which asymptotic response time per note is achieved: skilled group

Sequence length (notes)	1	2	3	4	5	
Weighting 1a	-4	1	1	1	1	$t(36) = 21.16, p < 0.0001$
Weighting 1b	0	-3	1	1	1	$t(36) = 9.34, p < 0.0001$
Weighting 1c	0	0	-2	1	1	$t(36) = 1.65, p = 0.11$
Weighting 1d	0	0	0	-1	1	$t(36) = 1.52, p = 0.14$

Table 10.4

Linear contrast analyses to determine the sequence length at which asymptotic response time per note is achieved: less skilled group

Sequence length (notes)	1	2	3	4	5	
Weighting 2a	-4	1	1	1	1	$t(40) = 17.48, p < 0.00001$
Weighting 2b	0	-3	1	1	1	$t(40) = 18.38, p < 0.00001$
Weighting 2c	0	0	-2	1	1	$t(40) = 2.30, p = 0.03$
Weighting 2d	0	0	0	-1	1	$t(40) = 0.60, p = 0.55$

10.2.4 One factor between (skill), two factor within (sequence length and structure) ANOVA

This analysis is only concerned with exploring the effect of structure, at sequence lengths of 3, 4 and 5 notes. As before, the data for the two levels of the between factor of skill are positively skewed, but once again the similar shapes and variances of the distributions (skilled SD = 163ms, less skilled SD = 217ms) make them acceptable for the purposes of ANOVA.

There is a highly significant main effect of structure ($F\ (1, 19) = 58.35, p < 0.0001$). With combined skill groups and sequence lengths, structured sequences exhibit a faster mean response per note than unstructured sequences: structured = 594 ms (SD = 199ms), unstructured = 665ms (SD = 213ms). The structure x sequence length interaction is not significant ($F_{2, 38}\ [0.04, 7.6] = 1.81, p = 0.18$), indicating that the effect of structure is consistent across sequence length levels. The interaction of skill x structure approaches significance ($F\ (1, 19) = 3.28, p = 0.086$), with the responses of less skilled participants showing a trend towards being more sensitive to structure than those of skilled participants. The presence of this trend suggests that the influence of structure should be explored for the individual skill groups. Considering that there is no statistical evidence of a significant structure x sequence length interaction, it would seem appropriate to investigate the role of structure using the combined levels of each skill group's sequence length data. The relevant data are presented in Table 10.5 and Figure 10.7. A strongly significant simple effect of structure is found for both skilled ($F\ (1, 10) = 35.71, p = 0.0001$) and less skilled participants ($F\ (1, 9) = 24.84$, $p = 0.0008$).

Table 10.5
Mean response time per note (ms) for each skill group at levels of structure (combined across sequence length, SD in parentheses)

	Skilled	Less skilled
Structured	522 (159)	666 (208)
Unstructured	576 (164)	754 (218)
Unstructured minus structured	54	88

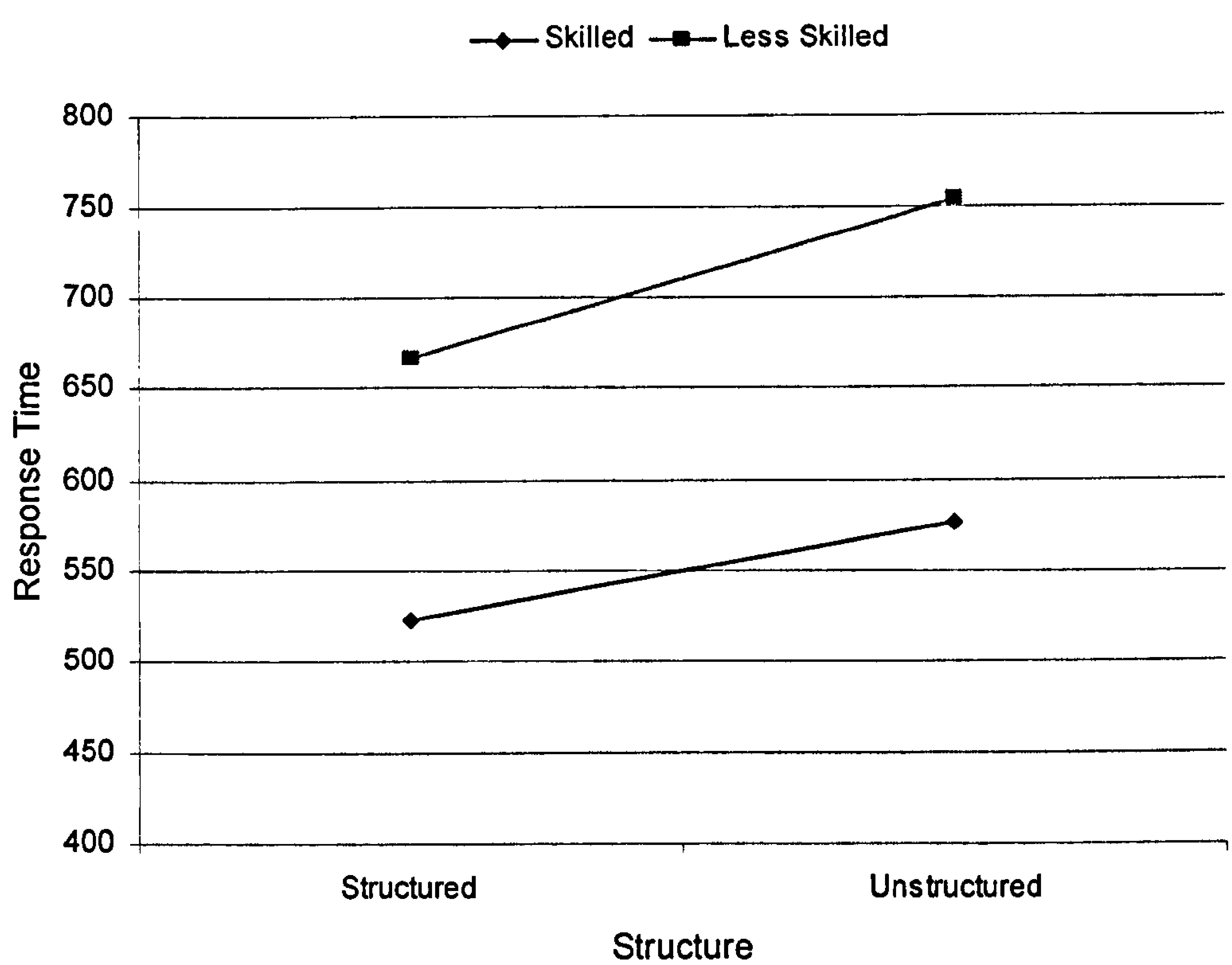


Figure 10.7
Mean response time per note (ms) for each skill group at levels of structure (combined across sequence length)

10.2.5 Other issues

Although using response times per note enables a generally more meaningful comparison to be made between results for different sequence lengths, one particularly interesting finding is more readily approached using complete response times. For each skill group, the mean scores obtained for the 2-note sequence length and for structured material at the 3-note sequence length are very similar. For the skilled readers the respective results are 1576ms and 1644ms, and for the less skilled readers, 1957ms and 2081. Post hoc t-tests carried out on these data segments indicate that there is no significant difference between these results for either group (skilled readers: $t = 1.10$, $p = 0.27$; less skilled readers: $t = 1.27$, $p = 0.20$). For unstructured material at the 3-note sequence length, the mean results are slower for both skill groups - 1806ms for skilled readers and 2450ms for less skilled. T-tests comparing these results with those of the 2-note sequence size reveal highly significant differences in the performance of both groups (skilled readers: $t = 3.57$, $p = 0.0004$; less-skilled readers: $t = 4.98$, $p < 0.0001$). Therefore, regardless of skill group, sequences of 3 notes may be processed as quickly as those of 2 notes, but only if they are clearly structured.

10.2.6 Error rates

Skilled readers performed, in total, 20 matching errors out of 880 tests (2.3 % error rate), and 13 key press errors. Less skilled readers performed 17 matching errors out of 968 tests (1.8 % error rate) and 20 key press errors. The error rate is therefore extremely low for both types of reader.

10.2.7 Summary of analysis

The skilled participants perform significantly and consistently faster than their less skilled counterparts across all sequence lengths, with no significant skill x sequence length interaction. Both groups of participants appear to reach asymptotic levels of performance at sequence lengths of 3 to 4 notes. Their average response times per note at these sequence lengths are approximately twice as fast as those achieved with single note sequences. Both skill groups display significant sensitivities to structure, obtaining consistently faster responses on structured material across the sequence lengths tested. The skill x structure interaction shows a trend towards significance, with less skilled participants showing evidence that their responses are more sensitive to the structural distinction than skilled participants.

10.3 Discussion

10.3.1 The effect of sequence length

The fact that the skilled group were found to reach their most efficient level of performance on the pattern-matching task at equivalent sequence lengths to the less skilled group is contrary to the expectations of the patterning account. According to that account, they should have achieved this within a larger notational context than less skilled readers. As with the controlled preview studies, this does not necessarily deny a primary role for structure in explaining the difference in performance between the skill groups in this study, but it does require, once more, a modified version of the patterning account in which skilled readers achieve their performance advantage from structural perception using an equivalent number of notes to less skilled readers. The next section will discuss whether the data offer any support for this modification. Before that, some consideration will be given to the meaning of the asymptotic response itself.

Without further evidence, for example relating to the eye-movements of subjects, it is not possible to draw any firm conclusions about the typical sizes of perceptual unit employed by the two skill groups in their comparison procedures. However, the fact that the most efficient level of performance is reached at sequence lengths of 3 or 4 notes for both skill groups suggests that where such a number of notes is available for perception, subjects make use of these unit sizes, rather than smaller ones. It is clearly not possible to conclude that these lengths mark the typical maximum size of unit used, although anecdotal evidence from post-experiment discussion with subjects about performance strategy does support the idea of the typical size being no more than 4 notes. Only two subjects, both from the skilled group, claimed they had made their 5-note sequence length comparisons on the basis of a single representation of the entire stimulus. Both of these had perfect pitch and stated that they had performed the trials predominantly using auditory representations.

The work of Truitt *et al.* (1997), already discussed in the literature review, does provide some support for the idea of 3 to 4 notes being the most efficient size of perceptual unit used by both skill groups in this study. They recorded identical effective perceptual span figures to these for both good and poor readers performing monophonic material at sight, suggesting that my findings may well have broader relevance to complete task performance, despite the different goals of the perceptual activity involved in the two tasks. It must be borne in mind though that the studies are not strictly comparable: Truitt *et al.* did not control for technical skill, and their perceptual spans specifically relate to a single fixation. Further issues are that we do not know the relative dependencies of the two measures on performance speed, or the extent to which either of them includes notes that have only been partially identified i.e. are effective in terms of their priming of subsequent perception. Without more research in these areas, therefore, it would be wise not to read too much into the similarity of the two sets of results. However, in the context of the discussion in Chapter 8 about the entire maximum effective preview span possibly not requiring explicit storage in short-term memory, 3 to 4-note perceptual units could well be sufficient to maintain the skilled maximum effective preview span of 6 notes recorded for monophonic material.

10.3.2 The effects of structure and skill

Both skill groups performed slightly faster with structured than with unstructured material: overall, skilled readers performed 9% faster and less-skilled readers 12% faster. These statistically significant differential responses to the structural distinction are of a similar size and trend to those for the two-part and four-part controlled preview studies (with the two-part study the equivalent performance gain for skilled and less skilled readers was 8% and 15% respectively, and with the four-part, 11% and 19%). It must be remembered though that the structural effect in these latter studies may relate to both local and larger scale structure, whereas with the current study it is confined to local structure. The fact that there was no equivalent significant effect of structure obtained for either skill group on the monophonic controlled preview study may indicate that other factors, possibly motor related, were limiting performance, preventing subjects from demonstrating their full perceptual capabilities.

As with the controlled preview studies, the data need to be examined to determine whether they support a primarily patterning-based explanation of skill difference - the modified patterning account - or whether they show skill difference to be principally the result of more elemental perceptual processing related factors. Considering the modified patterning account to begin with, an interpretation of the data discussed in relation to the two-part and four-part studies may again be appropriate. It relates to the fact that it is impossible to remove structure entirely from the unstructured sequences - it can only be disrupted. Because of this, there are circumstances in which the similar effect of structure obtained for each skill group might not actually represent an overall equivalent sensitivity to musical patterning. For example, it could be that a highly structurally sensitive skilled group were able to achieve close to maximum performance from their perception of available patterning within the unstructured sequences, with only a partial processing of cues within the structured sequences necessary in order to reach system performance limits. A certain amount of structure within the structured sequences would therefore have been surplus to the skilled readers' performance requirements. In contrast, the slower overall performance of the less skilled group could be put down to their being considerably less sensitive to patterning, their structural effect representing merely a rudimentary response to available cues. Such an interpretation requires that the skilled group were more sensitive to patterning within the unstructured sequences than the less skilled group were to patterning within the structured ones.

As with the controlled preview experiments, the appropriateness of such a patterning-based interpretation of the data is thrown into doubt by other results involving experimental conditions in which either no patterning, or only ambiguous patterning, was available for perception. There is no significant skill x sequence length interaction, meaning that the difference between the mean response time per note for the skill groups at the 1-note sequence length (187ms), where structure is irrelevant, and 2 notes (191ms), where structure is necessarily ambiguous, is very similar to that for longer sequences where the difference is more open to a structure-based interpretation because of the larger notational context (3 notes = 180ms; 4 notes = 141ms; 5 notes = 162ms). So, with skilled readers considerably faster at the task even at the 1-note sequence length level, there would seem to be no justification for invoking sensitivity to structure as an explanation for the effect of skill at larger sequence lengths. In the light of these results, the most plausible interpretation of the experimental data would therefore seem to be, firstly, that the effect of skill is principally the result of the skilled readers having faster elemental perceptual/transcriptional processing abilities in relation to the task than the less skilled readers and, secondly, that the similar responses of both skill groups to the structural distinction are indicative of an overall similar sensitivity to musical patterning. Further evidence for this latter point is provided by the fact that each skill group takes no longer to perform the task with three structured notes than they do with two notes, whereas they both take significantly longer to perform it with three unstructured notes (see section 10.2.5 above)

10.3.3 Summary and conclusion

The evidence from this study indicates that the skilled participants do not have a significantly different pattern of perceptual results from the less skilled participants, though they are faster overall at performing equivalent perception-related tasks. The finding that their performance advantage is not based upon the simultaneous processing of larger quantities of notation than less skilled participants is consistent with the results of the pattern matching experiments of Waters et *al.* (1997; 1998), and also Gilman and Underwood's (2003) error detection study that was discussed in the literature review. The finding is also in line with the general trend of results in eye-movement research discussed earlier in the thesis, research that fails to show any clear and consistent evidence of skilled readers using larger chunks than less skilled readers in their performance. Again, though, it is less in line with the typical interpretation of this research promoted by proponents of the patterning account, something that has been fully documented in the literature review. Concerning musical structure, the evidence that both skill groups have similar structural sensitivities in relation to perception is consistent with the limited previous perceptual work that has been carried out in this area (Waters *et al.* (1997), but clearly runs contrary to pattern account theorising. This evidence is fully in line with the influence of structure in my own controlled preview studies, however.

The above findings suggest that earlier concerns over the validity of the recorded measures as representations of sight-reading related perceptual ability were largely unfounded. This is further indicated by the strong correlation ($r_s = 0.76$, $p < 0.0001$) between the individual participant mean asymptotic responses on this study (results combined across structure and 3, 4 and 5-note sequence levels) and those of the monophonic preview study (measured as inter onset intervals, combined across structure and key levels). Overall, therefore, the pattern matching methodology would appear to have been an effective, sensitive tool for measuring perception-related ability in relation to the particular factors employed. Other interpretations of the results are possible, though. For example, the apparently simple distributional trends may be the result of a combination of more complex underlying causes. It is therefore important for this work to be repeated, preferably involving a study of associated eye-movements, something that might provide a more complete picture both of underlying processes involved in the task and of the typical chunk sizes employed by participants. It is also possible that no difference in the pattern of performance of the two skill groups was found because the materials used were too simple. Although Gilman and Underwood's dual stave study, just mentioned, would seem to question this, it would nonetheless be important to investigate a greater complexity of notational material, in terms of both parts and rhythm, to enable a broader picture of perceptual skill to be built up.

Chapter 11

Motor sub-skill study

11 Motor sub-skill study

This chapter is divided into three main sections:

- A description of the experimental methodology
- Results and statistical analyses
- Discussion of research findings

11.1 Methodology

11.1.1 Development of experimental design and technology

The rationale behind this study has already been briefly discussed in Chapter 3, but will now be expanded upon. As already mentioned in the literature review, without any particular empirical justification, cognitive research into sight-reading has traditionally tended to dismiss motor skill as an important factor in explaining variation in the sight-reading ability of skilled musicians. It seems to have been generally considered that both rehearsed and unrehearsed performance depends upon the same underlying motor mechanisms: 'the fact that musicians can have similar general performance abilities ('output' skills) but vastly different sight-reading abilities ('input + output' skills) implies that the attainment of input skills must be important to sight-reading facility' (Waters et *al.,* 1998a, p. 125). Ericsson and Lehmann (1996) specifically investigated motor skill and sight-reading ability, and proposed that skilled and less skilled sight-readers differed in an aspect of their rehearsed performance ability that became particularly exposed during sight-reading: the dependency of output upon visual monitoring. They argued that during sight-reading fairly continuous visual contact with the score needs to be maintained, but that rehearsed performance allows a greater compensatory viewing of the hands. They called the ability to perform without visual monitoring 'kinesthetic ability'.

In Ericsson and Lehmann's study, college level pianists sight-read, both with and without visual feedback, a partially rehearsed, short piano piece involving leaps across the keyboard. Skilled readers made fewer note errors than less skilled readers, all participants making more errors on the task where no visual monitoring was allowed. The researchers concluded that the results indicated the greater 'kinesthetic' ability of the skilled readers. However, this is not the only explanation; no consideration was given to the possibility that the skilled readers might have been capable of performing more accurately in the context of the limited amount of rehearsal. Thompson and Lehmann (2004) have since incorporated a distinction between rehearsed and unrehearsed movement into their theorising, emphasising the 'online' nature of the latter i.e. the need for novel motor programs to be devised at short notice. For example, they write that 'fingering choice may indeed be a determining factor of sight-reading expertise' (Thompson and Lehmann, 2004, p. 149). Such proposals would seem to be an implicit recognition of the inadequacy of the earlier research conclusions. The proposal of a distinction between skill at rehearsed and unrehearsed movement also requires the consideration that the two might vary in terms of their dependency upon visual monitoring. A further problem with Ericsson and Lehmann's study is that they did not sufficiently isolate the output of their participants from input skill variation meaning that the pattern of results obtained was not necessarily entirely motor related. Both skilled and less skilled participants played from standard music notation, the latter required to perform at the same speed as the former, something which may possibly have led to their experiencing a considerably greater input task load.

In the absence of clear evidence about the extent to which motor related factors play a significant role in explaining sight-reading skill variation, this issue was considered an important focus for this thesis. The approach chosen involved a development and refining of Ericsson and Lehmann's methodology. The proposed solution for achieving a greater isolation of motor ability from possible input skill variation was to dispense completely with standard notational input (something that will be elaborated upon later) and to require performance to be from memory. To facilitate this, trials consisted of only short sequences of notes, with all participants allowed sufficient time for secure memory encoding. Such a method also has the advantage that it prevents the visual tracking of input from interfering with output. As in the previous studies in this thesis, performance was self-paced at a maximum comfortable speed, enabling tempo to be used as a dependent variable in addition to errors. Since the aim was to focus on the most basic expressions of motor skill, experimental materials were chosen that lay under the hands, rather than involving performance leaps. Fingering was provided for participants in order to prevent fingering strategy from being a source of uncontrolled performance variation. Unlike Ericsson and Lehmann's study, participants were limited to purely mental rehearsal so that preparation for output would relate as closely as possible to unrehearsed sight-reading.

To draw conclusions about the degree to which performance variation is due to differing dependencies upon visual monitoring and to skill at unrehearsed movement, two experimental conditions are required: one involving no visual monitoring of the task and the other allowing complete visual monitoring of it. There was only sufficient experimental time to investigate one of these conditions, and non-visually monitored output was chosen because with both of the aforementioned factors possible sources of variation, it offered the greater chance of recording a performance difference between the two skill groups. If a difference was demonstrated, further research could then be undertaken to analyse its specific origins. Also, the controlled preview studies do provide data relating to partially visually monitored performance that may provide a helpful context for the interpretation of results.

To investigate motor output, the use of some form of input is unavoidable. The methodological challenge in relation to the current study was to find a simple means of input that avoided the use of standard notation, and that would be as neutral as possible in terms of its influence upon output. The preferred way of achieving this would have been to use an animation or video recording of the required hand movements on a piano keyboard, providing participants with a direct representation to store and prepare in memory, and then physically reproduce. The advantage of such an approach is the absence of any transcribed representations that might result in performance variation. It was also required that the technology incorporate the ability to measure the memorisation/preparation time of participants, and record the timing and pitch data of their performances. Once again, no commercial software was available that could carry out all the necessary tasks. Having to undertake the software development myself meant that some compromise had to be made in terms of experimental design. It was decided that a simpler, but still effective, approach would be to indicate the required movements by displaying numbers (representing standard piano fingering) upon selected keys of a graphical representation of a piano keyboard, the numbers cycling through one at a time to give the impression of movement. Clearly, the use of numbers meant that the need for some form of transcription was not entirely eliminated from this experiment. However, it was felt that the depiction of fingers as numbers would be familiar and intuitive for all participants. It must be borne in mind, though, that presenting stimulus in this manner may have led to mental representations untypical of normal sight-reading. Such a situation is not ideal, but the methodology does at least mark a step forward in the study of sight-reading motor performance. If it shows evidence of a useful approach, more sophisticated technology can be developed for future research. As was the case with previous software developed for this thesis, the program was written in the C language.

11.1.2 Materials

11.1.2.1 Basic design overview

No formal pilot study was carried out for this experiment, but as with the previous studies undertaken, the testing of the software with musicians during the development process amounted to an informal piloting. During this process the experimental method was fine-tuned and the most appropriate materials and condition levels chosen. As with the perceptual sub-skill study, the decision was made to concentrate entirely upon monophonic sequences of notes, so that a more complete understanding of foundational issues might be achieved within the limited experimental time available. To enable a direct comparison of results with those of the monophonic preview and perceptual sub-skill studies, performance materials of a similar nature to the ones used on those studies were designed. Sequences were again written within the range middle C to the G a 12th higher and were isochronous. Equivalent independent variable conditions to the perceptual sub-skill study were used. Again, the use of a condition of key was not appropriate given that the sequences were too short to enable any sense of home key to be defined. The melodic direction of sequences was not subject to the particular constraints of the pattern matching methodology, and so sequences were free to both ascend and descend as in the monophonic preview study.

11.1.2.2 Factors to be studied

Sequences were designed to investigate the influence of two factors on motor output: sequence length and structure.

Sequence length

A range of sequence lengths is useful in that it allows the investigation of whether skilled and less skilled readers have a different pattern of performance with increasingly long strings of output, something that might provide insight, for example, into how they vary in terms of capacity and security of short-term memory storage. Sequences were limited to 3 lengths: 3, 4 and 5 notes. 1 and 2-note sequence lengths were not used because they were considered to be overly simplistic. Sequences of 6 notes had been piloted during software development, but both skilled and less skilled readers typically found them very difficult to memorise within any reasonable length of time.

Structure

The sequences for each level of sequence length comprised three structural categories: structured, normal unstructured and difficult unstructured materials. Materials for the former two categories were designed to provide a structural distinction as equivalent as possible to that of the equivalent monophonic preview study, taking into account their considerably shorter lengths and the particular requirements of the experimental task (see below). The basic note collection for these sequences was the same as in the monophonic preview study – all the white keys together with F# and Bb. However, as with the pattern matching study, it was also necessary to include C# and Eb to enable a greater variety of unstructured material to be devised. For example, tritones were once more useful in designing these materials, but with only a limited number available within the diatonic scales of C, F and G there was the risk of overuse. Unlike the previous studies some limited chromatic use of black notes was made in the normal unstructured sequences to enable the greatest equivalence of movement between these and the structured materials. No more than one black note was used in either of these two types of sequences, so any resulting increase in complexity was negligible.

Difficult unstructured material required stretches over larger intervals, together with more awkward, non-standard fingerings, something that was facilitated by a greater use of all five black keys. The rationale for using such material was the concern, based upon experience during informal piloting, that the ordinary structured and unstructured sequences might have been too simple to differentiate between skilled and less skilled reader participants. This last category of sequence also made chromatic use of black notes.

Structured material

Structured materials for the pattern matching study sought to represent the most fundamental aspects of tonal melodic structure: close use of broken triads together with the employment of passing notes. To maintain the greatest structural integrity, all materials implied a single triad harmonically. Ideally, this would have been the preferred structural regime for the current study, but the use of passing notes presents a problem for research investigating motor output in the context of short note sequences: the domination of such sequences by three stepwise movements runs the risk of making them too simple to perform and therefore unable to distinguish between the abilities of skilled and less skilled participants. Not using passing notes in 3 or 4-note sequences would seem quite acceptable - it merely limits the range of melodic organisation that can be represented. However, with 5-note sequences, continuous broken chord movement starts to appear rather contrived especially when no change of harmony or more complex chord is implied. Also, such structure is not particularly representative of the more 'flowing' melodic organisation of the monophonic preview study materials. A problem with not using any stepwise movement at all in the structured materials is also that some movement of this nature is necessary to disrupt structure within the unstructured sequences, possibly making these latter sequences easier to perform.

There would appear to be no ideal solution to these matters and the compromise that was decided upon for the current study is as follows:

- 3-note sequences - based entirely upon typical, close broken chord movement using a single triad (major or minor)
- 4- note sequences - based upon typical, close broken chord movement using material from a single triad (major or minor), but with some sequences involving an appoggiatura to provide structurally appropriate stepwise movement.
- 5-note sequences - based upon typical, close triadic (major) and dominant 7th broken chord movement. Two triads/chords implying strong dominant to tonic or tonic to dominant progression, the chord change providing the opportunity for a single stepwise movement within these sequences.

All sequences were written to be as characteristic as possible of melodic movement within the monophonic structured sequences where passing notes were not used, given the constraints just discussed. Appropriate and comfortable fingering was provided for the sequences based upon how they might be performed within a larger notational context. In view of this, the required fingering did not always represent the most obvious, or indeed the easiest, strategy. The reason for this was again so that the materials should have the greatest chance of revealing any difference in skill amongst participants. Examples of these structured sequences are shown in Figure 11.1, and the complete collection is presented in Appendix 3.

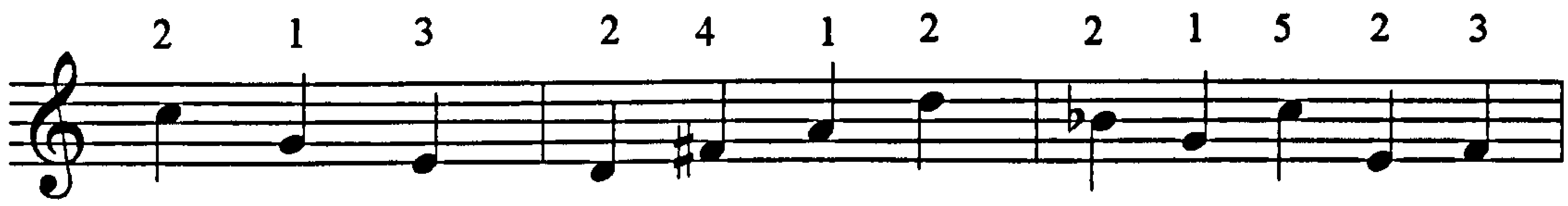


Figure 11.1
Examples of 3, 4 and 5-note structured sequences and fingering

Normal unstructured material

To ensure a reasonable consistency of melodic shape and interval size between the structured and unstructured sequences, each unstructured sequence was a degraded and transposed version of a structured sequence. Either two or three changes of triad (major or minor) were implied for the 3-note sequences, and at least three for the 4 and 5-notes sequences. Because 5-note structured sequences implied two triads, the structural distinction was possibly smaller with these materials. To try to ensure that performances with structured and unstructured material did not differ due to sequences varying in the difficulty of fingering movements, unstructured sequences were given broadly similar fingering movements to their structured equivalents. The number of sequences containing a black note was equivalent, within sequence length, to that of the structured materials. Examples of unstructured sequences are presented in Figure 11.2, and the complete collection can be found in Appendix 3.

Figure 11.2
Examples of 3, 4 and 5-note normal unstructured sequences and fingering

Difficult unstructured material

The category of difficult unstructured materials contained sequences that were equivalent in musical structural terms to the normal unstructured sequences. All sequences included a single black key but no more than two. The materials were either similar in intervallic spread to the normal unstructured sequences, in which case they were provided with more awkward fingering movements, or else involved larger intervals which allowed more difficult stretches to be included. Overall, the sequences were designed to lie significantly less comfortably under the hand than those of the other structure levels. This level of structure clearly does not allow any discerning of the relative extents to which structural perception and motor skill were influential upon performance. However, this was not the aim of this category of material – its role was primarily to provide participants with a high demand task that would afford the greatest opportunity of recording differences in monophonic motor skill. Examples of unstructured sequences are presented in Figure 11.3, and the complete collection can be found in Appendix 3.

Figure 11.3
Examples of 3, 4 and 5-note difficult unstructured sequences and fingering

11.1.2.3 Further design issues

In total, there were 54 individual sequences used in this study, constituted as follows:

1. 18 x 3 note tests. 6 structured, 6 unstructured, 6 difficult unstructured.
2. 18 x 4 note tests. 6 structured, 6 unstructured, 6 difficult unstructured.
3. 18 x 5 note tests. 6 structured, 6 unstructured, 6 difficult unstructured.

All the sequences were evaluated by the independent adjudicator. His task was primarily to confirm whether or not the sequences complied with the design brief described earlier, with regard both to structural content and fingering. However, he was also asked to consider whether any of the sequences, despite fulfilling the brief, might not be sufficiently representative of their 'structured' or 'unstructured' categorisation. For example, some unstructured sequences may have contained strongly implied structural elements that I had not been aware of. The adjudicator confirmed the categorisation of all materials.

11.1.3 Apparatus

The experiments were performed on the same Roland FP11 electronic piano used for the controlled preview experiments. The trials were again run out of the direct sight of participants, from a Pentium 4 laptop computer. Experimental materials were displayed on a 17-inch flat screen monitor situated directly behind the piano on an adjustable height table, approximately 80cm in front of participants. This monitor was linked to the external VGA port of the laptop, and the piano MIDI output was connected to the MIDI input interface of the laptop.

The screen interface of the software is shown in Figure 11.4. It consists of a graphical representation of a two-octave segment of piano keyboard commencing at middle C, presented at the centre of the computer screen, and set against a pale grey background. For individual trials, each of the keys to be performed was indicated by having the required performance finger number displayed upon it in red, making it stand out clearly from both white and black keys. The pitches of the required keys were not sounded so as not to give a particular advantage to participants with better aural ability. Each complete sequence of notes was cycled through in continuous manner until participants were ready to respond, one key highlighted at a time for 0.5 seconds and with a longer pause of 1.5 seconds between repetitions of the cycle. Once a piano key was depressed at the beginning of the response, finger numbers disappeared from the screen leaving participants entirely dependent upon stored representations. The computer recorded pitch and timing data for all keys played, the timer being initiated at the start of each sequence presentation.

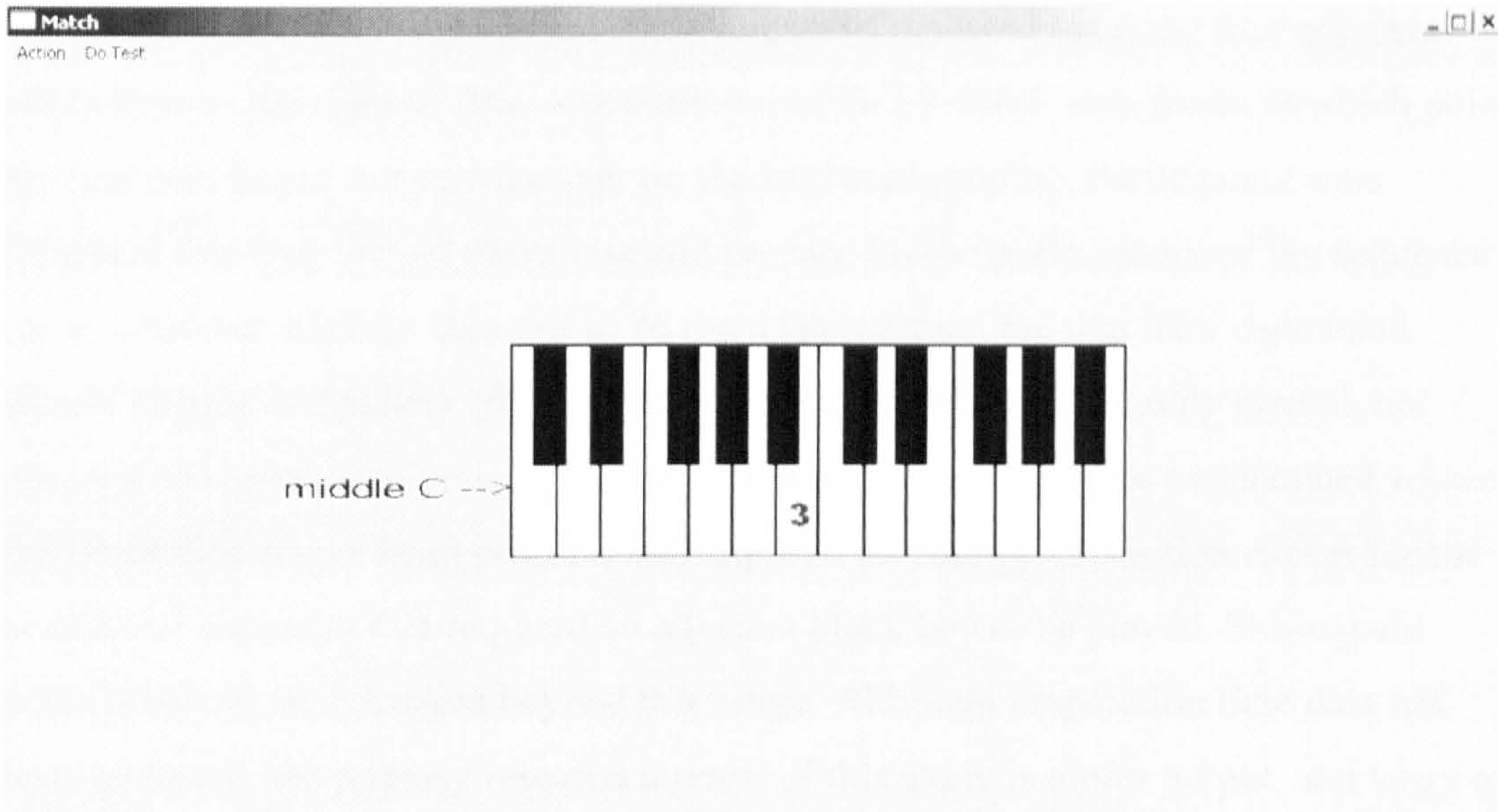


Figure 11.4
Screen interface of the motor sub-skill study software
(the middle of a sequence presentation indicating finger 3 to play B)

To prevent visual monitoring of performance, participants were required to wear a pair of safety goggles that had been adapted by adding a piece of card that projected horizontally forward for about nine inches from the base of the goggles, between the level of the nose and the eyes. The goggles were large enough to fit over spectacles. The keyboard was visible if the head was tipped forward significantly, something that participants were allowed to do to help position their hand in the correct location at the beginning of each new trial. In order that any inaccuracies of fingering could be identified, the entire series of trials was recorded on video, the video being positioned above and to the left or right of the piano (depending upon available room), focusing upon the hands only to ensure accurate transcription of the data. Participants were monitored in real time to ensure that they did not employ any large movements of the head during performance that might enable hands and keyboard to be seen.

11.1.4 Procedure

The same participants were used as for the previous experiments except that one member of the less skilled group did not take part because of illness. Prior to the formal experiment, participants were given as much practice at the procedure as they required in order to become comfortable with it. At the beginning of each trial, they were required to position their right thumb on a particular white key (which varied from trial to trial) and to place the other fingers of the hand upon the four adjacent white keys to the right of this. A countdown of '3-2-1-Start' was given, at which point the numbers began to cycle through on the keyboard graphic. Participants were informed that they should memorise and prepare for the performance of the sequence using whatever strategy they felt to be most appropriate, but that their right hand should remain completely still until they played the first note i.e. only mental, not physical rehearsal, was permitted. The first note of each sequence was situated within the initial five-finger hand position and required no change of position except for the occasional sequence that required an adjacent black key to be played. Subsequent notes involved an extension beyond this range. Although preparation time data has been analysed, the primary research interest of this study is motor output, and to try to ensure that responses were based upon secure representations, participants were encouraged not to rush the preparation process.

As with the previous experiments, participants were asked to perform the notes of the trials as quickly as comfortably possible, being mindful, however, that accuracy of notes and fingers should not be sacrificed for speed. If they were aware that they had pressed a wrong key, they were asked not to correct, but to continue their performance with an attempt at the subsequent material. If they did correct themselves, which occasionally occurred, the correction was not considered as part of their response data. If they could not remember a particular note or fingering, they were required nevertheless to complete the trial to the best of their ability. Each of the 10 skilled and 10 less skilled participants was presented with the 54 trials in a different random order, a short break being taken after the 18th and 36th trial. The entire experiment including explanation and training typically lasted for approximately 45 minutes. After the experiment, in a similar manner to the perceptual sub-skill study, participants were engaged in a short, informal discussion about the representational strategies they considered they had used during their preparations. The results of these discussions are analysed in Chapter 12.

11.2 Results

11.2.1 Introduction

There are 4 dependent variables:

- Preparation time per note. This measure was used rather than the total preparation time because it enables the efficiency of memorisation to be more easily compared for the different sequence lengths.
- Tempo - the mean value achieved for each sequence calculated from raw inter onset interval data. Within each sequence, the latter could not be considered sufficiently independent of each other to be used as data points for the analysis. Mean tempo was considered a more intuitive representation of the results than mean inter onset interval.
- Note errors - the performance of notes different from those indicated by the sequence presentation.
- Fingering errors - the performance of fingering different from that indicated by the sequence presentation (detected from video evidence).

The independent variables are:

- Skill (2 levels: skilled and less skilled reader groups)
- Sequence length (3 levels: 3, 4 and 5 notes)
- Structure (3 levels: structured, normal unstructured, difficult unstructured)

The analysis consists of three separate one-factor between subjects (skill), two-factor within subjects (sequence length and structure) repeated measures ANOVAs. The first ANOVA uses preparation time as the dependent variable, the second, mean tempo and the third, note errors. Fingering errors were relatively rare and therefore have not been subjected to statistical analysis.

11.2.2 Focal points of the analysis

The analysis of the preparation time component of this study is somewhat opportunistic. The measure was not designed to address any specific research question; it is merely incidental to the primary research focus of the study. Interpretation of the results relating to the measure is made difficult because they incorporate both memory encoding and mental rehearsal elements, and it is not possible to discern the durational proportion of either of these within the whole.

Despite this, they are potentially useful as an indicator of overall skill at the mental component of the task, and although the specific reasons for any variation may not be discernable, they have the potential to generate ideas for future research to consider.

The motor performance section of the analysis provides the potential for answering a number of important research questions about the role of motor ability in determining sight-reading skill. If skilled and less skilled readers have equivalent rehearsed performance skills, does it follow that they will have equivalent motor ability in relation to sight-reading, as seems to have been traditionally assumed by cognitive research within the domain? Or is there evidence that motor skills involved in sight-reading are different from those required by rehearsed performance, with skilled musicians varying in ability at them? The two aspects of sight-reading related motor activity measured in the current study are the ability to plan and execute physically unrehearsed movements, and the dependency of such movements upon visual monitoring. Unfortunately, as discussed earlier, the relative importance of these cannot be determined from the results because no corresponding data is available for fully visually monitored motor output. Finally, further questions that need to be addressed are whether the skilled readers have particular ability at performing longer sequences compared to less skilled readers (evidence of larger, more secure memory or more automated output mechanisms), and the extent to which skill at the experimental task is dependent upon the structural content of material performed.

11.2.3 Analysis of preparation time data

11.2.3.1 Introduction

Although the distributions of both skilled and less skilled participants are positively skewed, they are reasonably homogenous in terms of variance (skilled SD = 907ms, less skilled SD = 1609ms) and similar enough in shape to be considered acceptable for ANOVA. Where appropriate, the Greenhouse-Geisser adjustment has been applied to the degrees of freedom of individual effects in order to counter the lack of sphericity in the data, and such results have been displayed in the alternative form employed in the previous analyses. A standard significance level of 5% has been used. Lack of sphericity renders the standard multiple comparison tests unreliable, and so *post hoc* paired comparisons have been undertaken using t-tests, making the Bonferroni adjustment to the standard significance level i.e. dividing this level by the number of comparisons made. In this analysis 7 comparisons have been made, and therefore the significance level for these has been set at 0.7%.

11.2.3.2 Analysis

The analysis is organised under the headings of the three independent variable factors: skill, sequence length and structure.

Skill

The main effect of skill is significant ($F(1, 18) = 8.4$, $p = 0.01$), with the skilled readers demonstrating faster preparation times. The mean overall preparation time per note for skilled readers is 1741ms (SD = 907ms) and for less skilled readers 2705ms (SD = 1609ms).

Sequence length

The main effect of sequence length is highly significant ($F_{2,36}[1, 18] = 43.10$, $p < 0.0001$). For combined skill groups and structure levels, mean preparation time per note is 1807ms (SD = 924ms) for sequences of 3 notes, 2025 ms (SD = 1198ms) for sequences of 4 notes, and 2837ms (SD = 1716ms) for sequences of 5 notes. The principal source of this significant effect would appear to be the result for 5 notes, which requires considerably longer preparation time per note than for the other two levels. There is a significant interaction between the main effects of skill and sequence length ($F_{2,36}[1, 18] = 5.12$, $p = 0.036$), and the relevant data are presented in Table 11.1 and Figure 11.5. These data demonstrate the same general trend for both skill

Table 11.1

Mean preparation time per note (ms) for each skill group at levels of sequence length (combined across structure, SD in parentheses)

Sequence length (notes)	3	4	5
Skilled	1504 (607)	1558 (757)	2161 (1125)
Less skilled	2110 (1076)	2493 (1366)	3514 (1928)
Less skilled minus skilled	606	935	1353

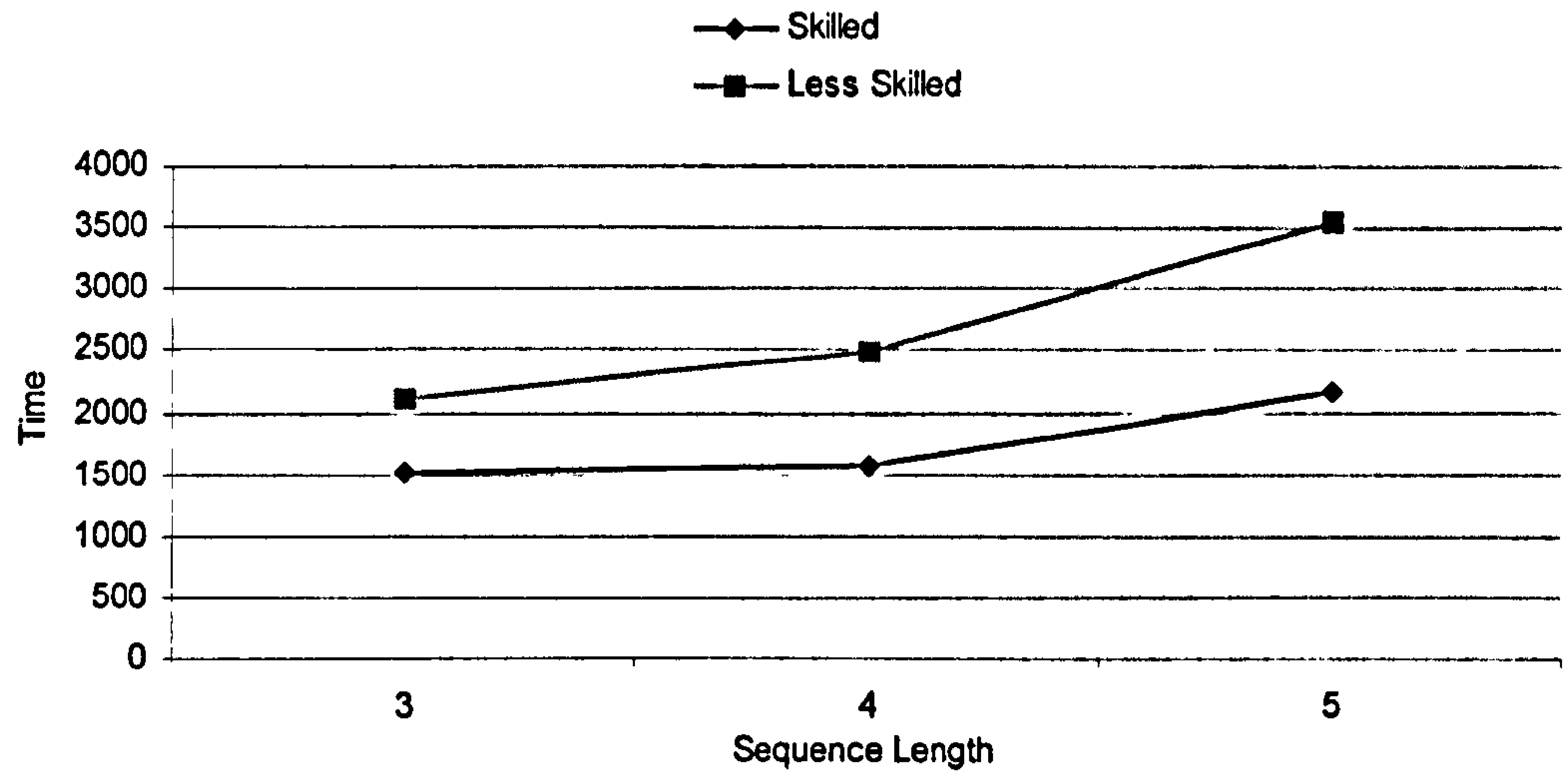


Figure 11.5
Mean preparation time per note for each skill group (ms) at levels of sequence length (combined across structure)

groups as for the combined results just described, but with the less skilled group performing more slowly at each sequence length and their performance increasingly lagging behind that of the skilled group with increasing sequence length. A *post hoc* t-test reveals that even the smallest of the differences (at 3 notes) is highly significant ($t = 6.37$, $p < 0.0001$). The significant interaction is found to relate principally to 3 and 4-note sequence lengths, evidenced by the fact that simple interaction at 3 and 4-notes is significant ($F(1, 18) = 10.35$, $p = 0.005$), but that at 4 and 5-notes it is not ($F(1, 18) = 2.91$, $p = 0.1$). There is a significant simple effect of sequence length both for the skilled group ($F_{2,18}[1, 9] = 14.49$, $p = 0.004$) and for the less skilled group ($F_{2,18}[1, 9] = 28.97$, $p < 0.0004$). *Post hoc* t-tests have been carried out to compare results within each skill group between the 3 and 4-note, and between the 4 and 5-note sequence length levels. Between 3 and 4 notes, a significant difference is found for the less skilled group ($t = 3.27$, $p = 0.001$) but not for the skilled group ($t = 0.97$, $p = 0.33$). Between 4 and 5 notes both skill groups show a significant difference in response (skilled: $t = 6.95$ $p < 0.0001$; less-skilled: $t = 7.31$, $p < 0.0001$).

<u>*Structure*</u>

The data demonstrate a highly significant main effect of structure ($F_{2,36}[1, 18] = 22.62$, $p = 0.0002$). Mean preparation times per note for combined skill groups and sequence lengths are 2016ms (SD = 1155ms) for structured material, 2163ms (SD = 1365ms) for normal unstructured material, and 2490ms (SD = 1584ms) for difficult unstructured material. These results indicate that the principal source of the main effect is the response to difficult unstructured sequences. The interaction of skill x structure (F (2, 36) = 2.83, p = 0.11) is non-significant. *Post hoc* t-tests show that the difference between the combined results of both skill groups at the two unstructured levels is significant (t = 3.97, p < 0.0001), but that the equivalent difference at structured and normal unstructured levels is non-significant (t = 2.16, p = 0.03).

The structure x sequence length interaction is just significant ($F_{4, 72}$ [0.5, 9] = 6.23, p = 0.05) and Table 11.2 and Figure 11.6 present the data for this. The principal source of the interaction would appear to be the data for structured material, where the response at the 4-note sequence length appears to be unusually fast. It is difficult to account for this, but one explanation might be that in the context of this particular task, the structural distinction is most effective with 4 notes, 3 notes perhaps being too simplistic for the perception of structure to influence preparation time, and 5 notes too demanding. The skill x structure x sequence length interaction is non-significant ($F_{4, 72}$ [0.5, 9] = 3.31, p = 0.11) indicating that the data for both the skill groups follow a similar pattern to their combined representation in Figure 11.6.

Table 11.2
Mean preparation time per note (ms) for levels of sequence length at levels of structure (combined skill groups, SD in parentheses)

Sequence length (notes)	3	4	5
Structured	1788 (793)	1623 (813)	2636 (1462)
Normal Unstructured	1629 (823)	2117 (1306)	2743 (1615)
Difficult Unstructured	2003 (1094)	2336 (1302)	3133 (1999)

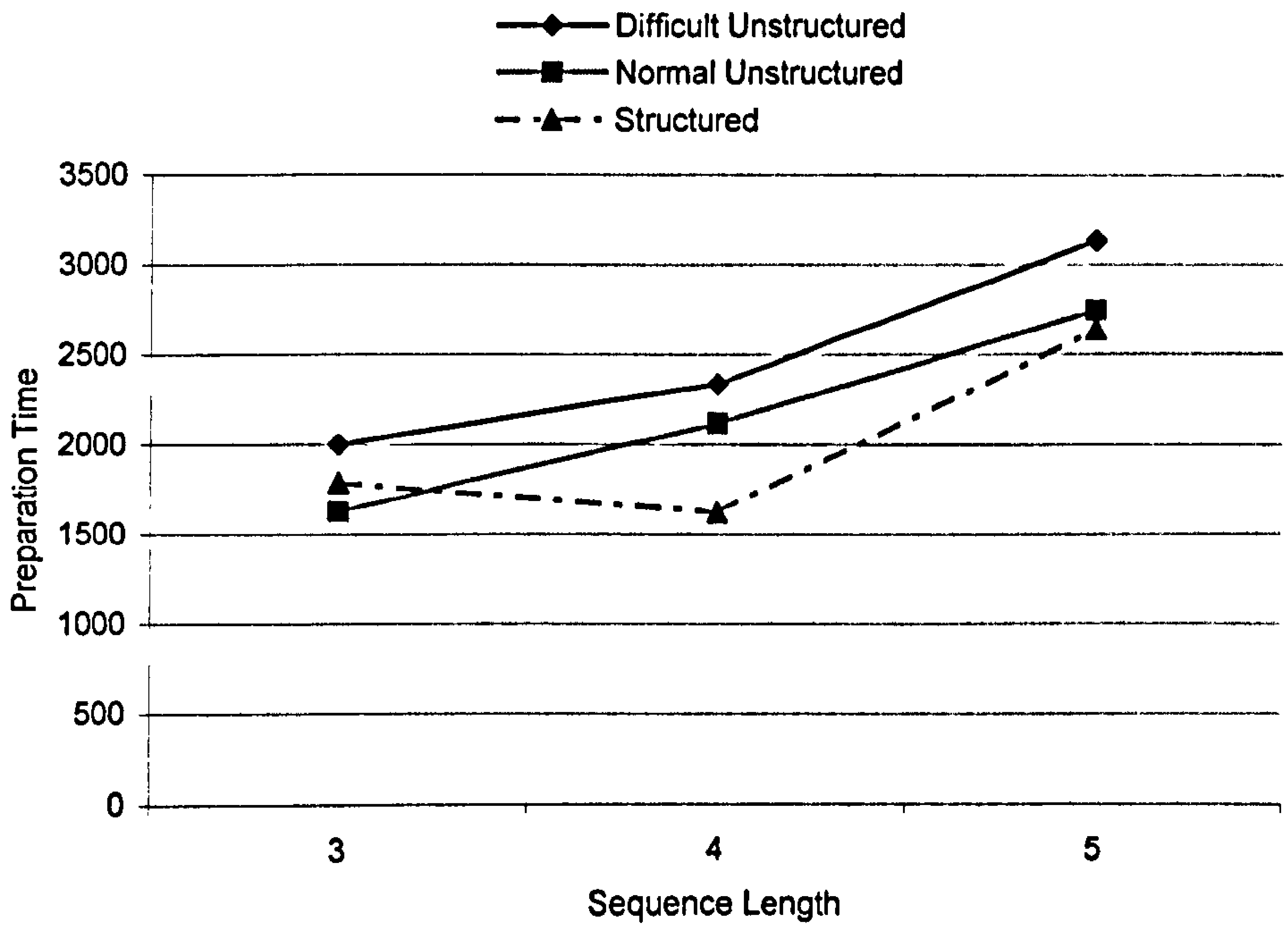


Figure 11.6
Mean preparation time per note (ms) for levels of sequence length at levels of structure (combined skill groups)

11.2.4 Analysis of tempo data

11.2.4.1 Introduction

The distributions of data for both the skilled and less skilled participants are again positively skewed. As previously though, given their similar shapes and variances (skilled SD = 171ms, less skilled SD = 367ms), they are acceptable for ANOVA. Greenhouse-Geisser adjustments have again been made to counter the lack of sphericity. A standard significance level of 5% has been used, and as with the preparation time analysis, *post hoc* comparisons have been undertaken with t-tests, using the Bonferroni adjustment to the standard significance level. In this analysis there are five such comparisons, making the required significance level 1%.

11.2.4.2 Analysis

The main effect of skill is highly significant ($F(1, 18) = 16.51$, $p = 0.0007$). The overall mean tempo for skilled readers is 128 notes per minute (SD = 46), and for less skilled readers 86 notes per minute (SD = 29). The effect of structure is also highly significant ($F_{2,36}[1, 18] = 42.33$, $p < 0.0001$). With combined skill groups and sequence lengths, the mean tempo is 119 notes per minute (SD = 45) for structured material, 114 notes per minute (SD = 49) for normal unstructured material, and 88 notes per minute (SD = 30) for difficult unstructured material. The main effect of sequence length is non-significant ($F (2, 36) = 1.89$, $p = 0.26$) as is the skill x sequence length interaction ($F_{2, 36} [1, 18] = 0.16$, $p = 0.85$), indicating that neither skill group responds in any significant manner to sequence length.

The skill x structure interaction is bordering on significance ($F_{2, 36} [1, 18] = 3.87$, $p = 0.065$) and therefore warrants further investigation. The relevant results for the individual skill groups are presented in Table 11.3 and Figure 11.7, and demonstrate a trend similar to the combined data given in the last paragraph - similar tempi at structured and normal unstructured levels, and considerably slower performance at the difficult unstructured level. On average, the less skilled participants always perform more slowly than the skilled participants, a *post hoc* t-test indicating that even the smallest performance difference between the skill groups (at the difficult unstructured level) is highly significant ($t = 9.54$, $p < 0.0001$). The simple skill x structure interaction at the two unstructured levels is significant ($F (1, 18) = 5.05$, $p = 0.04$) but the equivalent simple interaction with structured and normal unstructured material is non-significant ($F (1, 18) = 0.58$, $p = 0.46$). This indicates that the principal source of the main interaction lies in responses to the difficult unstructured material. Both skill groups demonstrate significant simple effects of structure (skilled: ($F_{2, 18} [1, 9] = 21.39$, $p < 0.0001$; less skilled: ($F_{2, 18} [1, 9] = 23.68$, $p = 0.001$. *Post hoc* t-tests reveal a highly significant difference between performances at the two unstructured levels for

Table 11.3
Mean tempo (notes per minute) for each skill group at levels of structure (combined across sequence length, SD in parentheses)

	Structured	Normal unstructured	Difficult unstructured
Skilled	142 (46)	139 (52)	104 (26)
Less skilled	96 (29)	90 (28)	73 (25)
Less skilled minus skilled	46	49	31

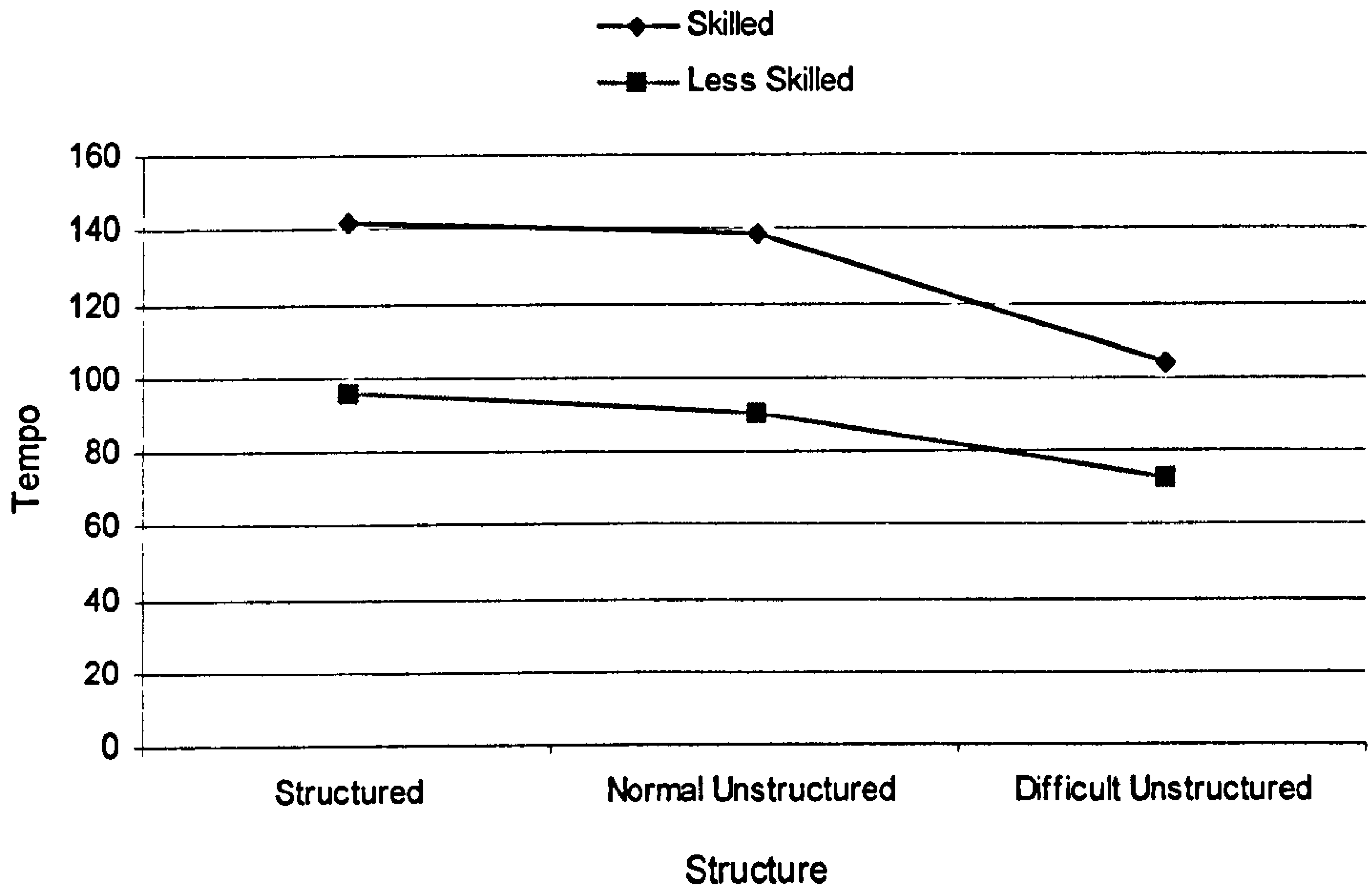


Figure 11.7
Mean tempo (notes per minute) for each skill group at levels of structure (combined across sequence length)

both skilled readers ($t = 9.22$, $p < 0.0001$) and less skilled readers ($t = 7.79$, $p < 0.0001$). Between the structured and the normal unstructured levels there is a significant difference for less skilled readers ($t = 2.77$, $p = 0.006$) but not for skilled readers ($t = 1.69$, $p = 0.09$).

The structure x sequence length interaction also borders on significance ($F_{4, 72}$ [0.5, 9] = 5.3, $p = 0.06$) and therefore requires some investigation. The relevant data are displayed in Table 11.4 and Figure 11.8. It would appear to be the result for unstructured material at the 3-note sequence length that is incongruous, performance being faster than with structured material. As in the preparation time analysis, it is difficult to draw any firm conclusions about this. One possibility is that the task demand of structured and normal unstructured sequences of 3 notes is simple enough for performance to be maximised regardless of structural content. The non-significant skill x sequence length x structure interaction ($F_{4, 72}$ [0.5, 9] = 0.42, $p = 0.40$) indicates that both skill groups demonstrate similar patterns of response to the combined distributions seen in Figure 11.8.

Table 11.4
Mean tempo (notes per minute) for levels of sequence length at levels of structure (combined skill groups, SD in parentheses)

Sequence length (notes)	3	4	5
Structured	119 (43)	123 (47)	115 (46)
Normal unstructured	122 (62)	111 (41)	109 (39)
Difficult unstructured	88 (28)	86 (30)	92 (32)

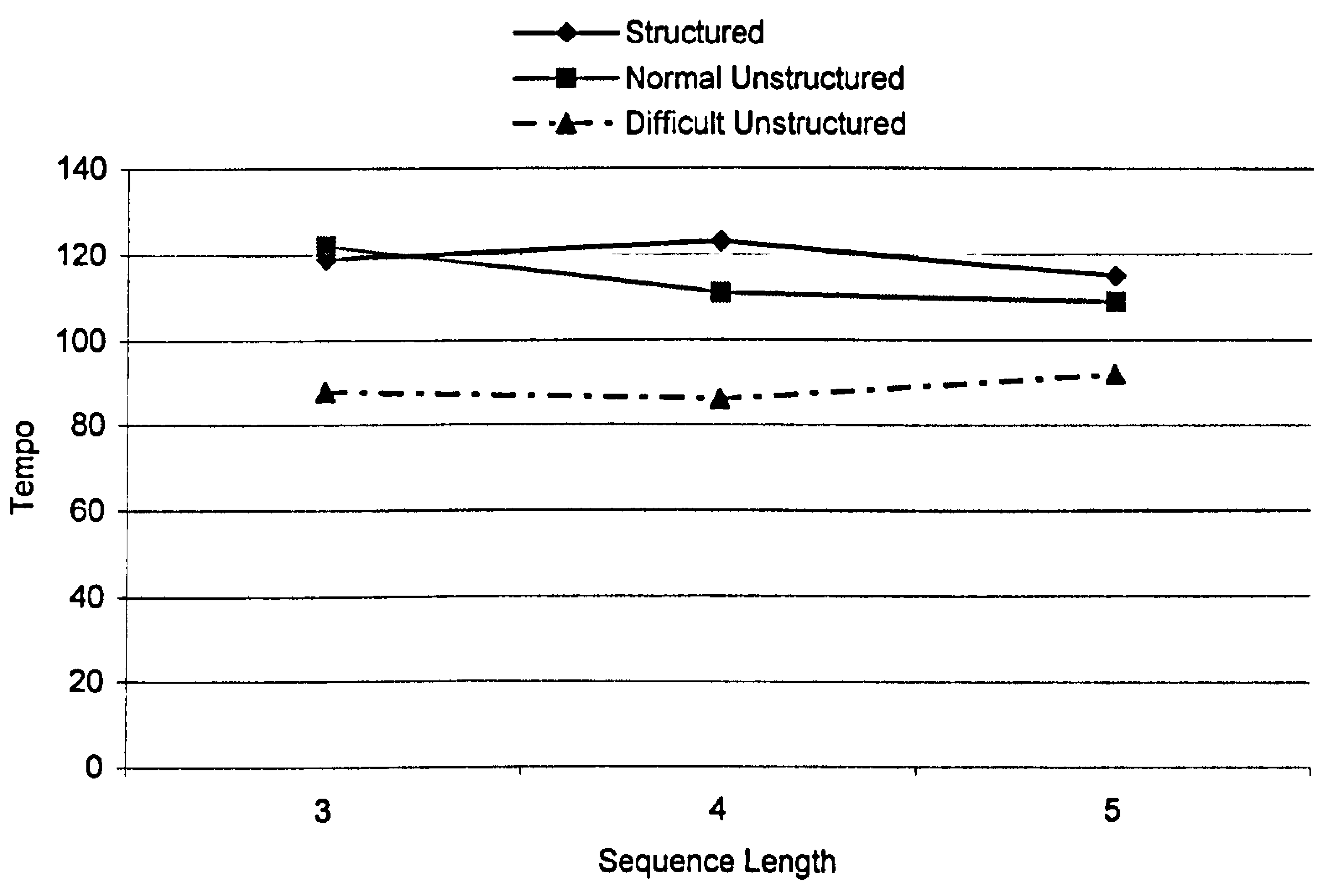


Figure 11.8
Mean tempo (notes per minute) for levels of sequence length at levels of structure (combined skill groups)

11.2.5 Analysis of errors

There were a total of 216 notes to be performed by each participant. The average number of errors per participant for the skilled group was 11.7 note errors (5.3% error rate) and 3 fingering errors (1.6% error rate). The equivalent figures for the less-skilled group were 26 note errors (12% error rate) and 11 fingering errors (5.1% error rate). The note error rates are considerably larger than the error rate in the other studies and would therefore be worth examining in more detail.

A 1-factor between subjects (skill), 2-factor within subjects (sequence length and structure) repeated measures ANOVA carried out on the percentage error rate data shows significant effects of skill ($F(1, 36) = 8.4$, $p = 0.01$), structure ($F_{2,36}[1, 18] = 13.62$, $p = 0.002$) and sequence length ($F_{2,36}[1, 18] = 8.92$, $p = 0.008$). There are no significant interactions. Five *post hoc* t-tests have been carried out on the combined skill group data to further investigate the data, using a significance level modified by the Bonferroni adjustment (5% divided by 5 = 1%). From these tests it would appear that it is the data at the difficult unstructured level that is the source of the significant effect of structure. The difference between the combined results of the skill groups for structured and normal unstructured material (combined sequence lengths) is not significant ($t = 0.053$, $p = 0.96$) but the equivalent difference between the results for normal unstructured and difficult unstructured material is significant ($t = 3.85$, $p = 0.002$). The relevant data are presented in Table 11.5. It also seems that the significant effect of sequence length is principally driven by the difference between the 3-note and 5-note sequence data. The difference between the combined results of the skill groups for these two levels (combined structure levels) is significant ($t = 4.45$, $p < 0.0001$), but the equivalent results for 3 notes and 4 notes, and 4 notes and 5 notes are not, although they do approach significance (for the former, $t = 1.82$, $p = 0.07$; for the latter, $t = 2.07$, $p = 0.03$). The relevant data are displayed in Table 11.6.

Table 11.5

Percentage note error rate for combined skill groups at levels of structure (combined across sequence length)

Structure category	Structured	Normal unstructured	Difficult unstructured
Combined skill groups	6.5%	6.6%	11.8%

Table 11.6

Percentage note error rate for combined skill groups at levels of sequence length (combined across structure)

Sequence length (notes)	3	4	5
Combined skill groups	5.8%	8.1%	10.9%

Table 11.7 presents data relating to how far, in semitones, individual note errors were from the note that was actually required. For errors of 2 or 3 semitones in size, both skill groups score very similarly. Skilled readers however perform relatively more errors of 1 semitone, whereas less skilled score relatively more errors of 4 semitones and larger. However, such differences are not really large enough to indicate any significantly different pattern of behaviour.

Table 11.7

Note errors for each skill group (as percentage of total note errors) categorised according to distance from required note (in semitones)

Distance (semitones)	1	2	3	4-plus
Skilled	25%	58%	14%	3%
Less skilled	18%	57%	15%	10%

Considering the low fingering error rate, particularly for skilled participants, statistical analysis would seem less appropriate for these data. To enable a basic comparison to be made, Table 11.8 presents the data for each skill group at levels of sequence length (combined structure levels) and Table 11.8 for levels of structure (combined sequence length levels). No clear trend is apparent for skilled participants, but less skilled participants appear considerably more prone to fingering errors at sequence lengths of 5 notes and also with difficult unstructured material. Indeed 46% of all their fingering errors occurred at the former level, and 67% at the latter. A final point worth mentioning regarding this category of errors is that for both skill groups, nearly half of their total errors occurred on the last note of a sequence (skilled = 46%, less skilled = 45%).

Table 11.8

Percentage fingering error rate for each skill group at levels of sequence length (combined across structure)

Sequence length (notes)	3	4	5
Skilled	1.1%	2.6%	1.1%
Less skilled	2.0%	3.5%	8.2%

Table 11.9

Percentage fingering error rate for each skill group at levels of structure (combined across sequence length)

Structure category	Structured	Normal unstructured	Difficult unstructured
Skilled	1.4%	1.7%	1.8%
Less skilled	4.0%	4.2%	7.1%

11.2.6 Summary of results

The less skilled group take significantly longer than the skilled group to encode and mentally rehearse material at each level of structure and sequence length. There is no significant difference between the skilled participants' efficiency of preparation at 3 and 4-note sequence lengths, but they do become significantly less efficient with 5 notes. Less skilled participants demonstrate a significantly reduced efficiency in preparation from both 3 to 4-note and from 4 to 5-note sequence lengths. The skill groups do not differ significantly in their pattern of response to structure. Both groups are significantly slower with difficult unstructured material than with normal unstructured material but show no significant difference in response between structured and normal unstructured material. There is no evidence for either skill group that the effect of structure varies with sequence length.

The skilled group perform the sequences at significantly faster tempi than the less skilled group across all structure levels and sequence lengths, with a greater level of accuracy. There is no significant effect of sequence length for either skill group. Concerning structure, skilled participants show no significant difference in tempo between structured and normal unstructured materials, but a significant one between normal unstructured and difficult unstructured materials, performing more slowly on the latter. Less skilled participants perform significantly more slowly with normal unstructured material than with structured material, and also with difficult unstructured material compared to normal unstructured material. There is no evidence that these patterns of data vary for different sequence lengths.

The evidence indicates that less skilled participants make significantly more note errors than skilled participants, but that there is no difference in the overall pattern of their error behaviour. For both groups, the note error rate increases with sequence length, but with regard to structure, it is only the performance of difficult unstructured material that leads to a significant increase in error rate. The fingering error rate is considerably lower for both skill groups than the note error rate. This rate would not appear to vary for skilled readers in relation to structure and sequence length, but for less skilled readers the same trend in results is apparent as for note errors.

11.3 Discussion

11.3.1 Preparation Time

11.3.1.1 The effects of skill and structure

With 3-note sequences, where both skill groups show the greatest efficiency of preparation, the less skilled group take on average 40% longer per note than the skilled group in encoding and planning their performances (skilled = 1504ms per note, less skilled = 2110ms per note (combined structure levels)). This result is similar to the findings of the perceptual sub-skill study where, at asymptotic levels of efficiency, less skilled participants were 45% slower in their responses than skilled participants. However, perhaps not too much should be read into this similarity. Whilst they may indeed indicate that less skilled participants are considerably slower than skilled participants at encoding monophonic material even when no transcription from standard notation is involved, it is also possible that skilled and less skilled participants may have encoded at similar rates, and that the source of the latter group's slower response lies perhaps in weaker output planning mechanisms, requiring longer mental rehearsal times.

The absence of any significant difference in performance between structured and normal unstructured material for either skill group is consistent with the findings of the other study involving monophonic output, the monophonic preview study, but contrasts with the significant effect of structure found in the similarly monophonic perceptual sub-skill study. The lack of a significant structure x sequence length interaction is also in line with the perceptual sub-skill study findings, and with the lack of a structure x preview size interaction in the monophonic preview study. As with the latter study, the absence of a structural effect in the context of materials of equivalent complexity may have a number of possible explanations. It could be that the task was too simple for any structural distinction to be useful to participants, with the results indicative therefore of raw encoding and/or motor planning skill. Alternatively, the task may have been demanding enough for a structural effect to register, but the structural distinction may not have been sufficiently clear. In such a case, the data may be consistent with a patterning-based account, with skilled participants finding sufficient structure even within the normal unstructured materials to maximise performance, and with the slower response of less skilled participants on both types of material a result of their insensitivity to patterning. Another possibility is that the extra demand of storing fingering as well as notes may have cancelled out any performance advantage gained from structure itself. Of course, it may also be that the novelty of the stimulus presentation simply caused participants to be less attentive to structural content. Whilst all these explanations are theoretically possible, the patterning account would again seem questionable, for similar reasons as before: although it is conceivable that highly patterning-sensitive skilled readers may have found sufficient structure within the context of unstructured 5-note sequences to enable their speed of encoding to match that of structured material, that the same could have occurred with structurally impoverished 3-note sequences seems implausible.

11.3.1.2 The effect of sequence length

The data indicate that both participant groups are able to successfully encode up to the maximum size of sequence length used - 5 notes. Obviously, no conclusion can be drawn as to whether skilled readers have larger short-term memory capacity in relation to the task than less skilled readers in view of there being no explicit data for longer sequences. However, anecdotal evidence suggests that the two groups may have very similar storage capabilities. As previously discussed, 6-note sequences were piloted, but not used for the actual study because participants of both skill types found them very difficult to memorise within any reasonable time frame. Although there is no clear evidence relating specifically to absolute short-term memory capacity, skilled participants do become significantly less efficient at the preparation process beyond 4-note sequences, whereas less skilled participants become less efficient beyond 3-note sequences. If this difference in the point at which preparation efficiency starts to decline significantly relates specifically to encoding, rather than mental rehearsal activity, it may provide skilled participants with what is effectively a short-term memory capacity advantage over less skilled readers when considered in the context of a time-constrained activity like sight-reading. In other words, it may not be the actual capacity of short-term memory storage that is limiting to less skilled readers' performance for more complex musical materials, as suggested by the four-part preview study, but rather the availability of memory that is 'fast' enough to sustain further increases in performance speed. Although the findings of the current study relate to monophonic materials, there was no evidence in the monophonic preview study of any such short-term memory size effects influencing skill difference. This apparent contradiction is probably explained by the demands upon memory resources being considerably greater in the sub-skill study, with its requirement for explicit memorisation of both notes and imposed fingering.

11.3.2 Motor performance

11.3.2.1 Introduction

The results show that the less skilled group are on average significantly slower and more error prone than the skilled group at performing the visually unaided, unrehearsed motor tasks in this study. The evidence would seem to indicate that the difference in output between the groups is primarily motor-driven i.e. that the study has been quite effective in isolating output from input related variation. First of all, there would seem to be no sign that faulty encoding or less secure memory storage was significantly influential in the greater number of errors performed by the less skilled participants. If this were the case, one might have expected the note error rate of less skilled participants to increase relative to that of the skilled participants with longer sequences, but no significant interaction was present here. Also, error notes were typically very close to the target notes, suggesting problems with control of execution rather than encoding and storage. For less skilled participants, 90% of errors fell within three semitones of the required note, the equivalent figure for skilled participants being 97% (see Table 11.7). This explanation is further supported by the common occurrence that where there was a run of two or more error notes, these often represented a correct transposition of the required target notes. In such cases, participants may have been unaware that they had played the initial note incorrectly, or if they were aware, perhaps found it difficult to alter their already planned output for the subsequent note. The data for fingering errors, however, does provide some evidence that security of storage may not be entirely irrelevant to explaining the difference in output of the skill groups, although it must be borne in mind that the proportion of notes fingered wrongly was very small. At 5-note sequence lengths, less skilled participants performed a considerably larger number of fingering errors than skilled participants, compared to the results for shorter sequences. Evidence that these errors are memory related, rather than simply oversights, might be provided by the large proportion of errors (45%) in which the last note of a sequence was fingered wrongly. Interestingly, the skill groups scarcely differ in terms of the proportion of fingering errors that fall on the last note of sequences (skilled group = 46%), indicating that the skilled readers were prone to similar lapses.

Secondly, with regard to the slower performance tempi of less skilled participants, the fact that there is no significant effect of sequence length would suggest that even the longest sequences were typically based upon a secure encoding of the materials i.e. there are no signs that less skilled participants were less efficient at retrieving longer sequences from memory. Evidence specifically supportive of an output-based explanation for the less skilled group's overall slower tempi is provided by the video documentation. Unlike skilled readers, and also unlike their own asymptotic monophonic preview performance, many of the less skilled readers often resorted to feeling their way across the surface of the keyboard to find the location of the next key, a strategy that was clearly responsible for at least some of their lower rate of output and that implied a particular deficiency with regard to geographical/spatial representations of keyboard layout.

11.3.2.2 The effect of skill

Both groups perform at considerably slower tempi than those achieved at maximum effective levels of preview on the monophonic preview study, therefore significantly understating their actual sight-reading related motor capabilities. The skilled group perform on average at 128 notes per minute compared to a mean maximum speed of 299 notes per minute on the monophonic study; for less skilled readers, the equivalent figures are 86 and 204 notes per minute. Skilled participants' results lie between their 1-beat and 2-beat preview performance levels (105 and 174 beats per minute), but less skilled participants on average do not even reach their mean 1-beat preview tempo (96 beats per minute. These data might suggest that participants were organising their performance of sequences on a note-by-note basis rather than as fluently planned, integrated motor programs. This idea receives some confirmation from the video documentation that has already been described in relation to less skilled participants. Although skilled participants did not resort to such tactile manoeuvring, their movements were noticeably more cautious and disjointed, contrasting with their fluent asymptotic performance on the monophonic preview study. The decline in motor performance from asymptotic monophonic preview study levels is no doubt due to the blindfold element of the task, and would seem to confirm the finding of Banton (1995) and Ericsson and Lehmann (1996) that some visual monitoring, perhaps only involving peripheral vision, is necessary for efficient sight-reading output. Nonetheless, comments made by participants after performing the experiment, suggest that skilled participants, in particular, might have been capable of considerably faster output, but tended to play safe in order to be as sure as accurate as possible. Refinements to the methodology are clearly necessary to ensure that results represent a maximising of participants' performance. One way to achieve this would be to pace output at a variety of different speeds, a technique that would enable the fastest speed for accurate performance to be determined.

Although, the experimental data may not represent maximum levels of performance, it would nonetheless seem reasonable to conclude that the variation in output between the two skill groups on this study is relevant to explaining the difference in their sight-reading ability. Not only is the variation statistically significant, there is also a strong correlation ($r_s = 0.67$, $p = 0.001$) between the individual participant mean tempi (combined across structure and sequence length levels) and those of the monophonic preview experiment at maximum effective preview span levels and above (combined across structure and key levels). As discussed earlier, to determine the extent to which the variation is the result of differences in unrehearsed motor ability and of differing dependencies upon visual feedback, further research now needs to be carried out with visually monitored output. It would also be important to find a way to control the degree of visual monitoring available to participants to enable its role within sight-reading related motor activity to be more fully understood.

The prevention of visual feedback did not lead to the less skilled readers performing proportionally more note errors than the skilled readers compared to the controlled preview studies. The less skilled/skilled error ratio for the motor sub-skill study was 2.3 to 1, and the mean equivalent ratio for the three preview studies was 3.1 to 1. It clearly cannot be concluded from this that the skill groups are similarly dependent upon visual monitoring of output, however, because there is no data concerning the extent to which their complete task performances were visually aided. The less skilled/skilled error ratio on the pattern matching study was 0.8 to 1, something that suggests that the greater error rate of less skilled participants in the studies involving a motor component may be related more to output rather than perceptual or cognitive processing factors.

<u>11.3.2.3 The effects of sequence length and structure</u>

In view of the impoverished levels of output achieved by both skill groups, the data for the sequence length and structure conditions are unable to provide any valid insight into unrehearsed motor activity within a normal sight-reading context. The further empirical work, just described, needs to be undertaken before the research questions posed earlier in the chapter can be addressed. Despite this, however, there are some related issues to consider.

First, there is no significant difference in performance speed for skilled participants between the structured and normal unstructured condition (structured = 142 notes per minute, normal unstructured = 139 notes per minute (combined sequence length levels)). With less skilled participants the difference is significant, but only small (structured = 96 notes per minute, normal unstructured = 90 notes per minute (combined sequence length levels)). This overall result is in line with the general trend of the other studies, which have typically shown the less skilled participants to have relatively larger responses to structural distinctions than skilled participants. However, the effect is only a small one, here, and so perhaps not too much should be read into it. Second, the difference in performance between the skill groups declines considerably at the difficult unstructured level – it is 46 notes per minute at the structured level, 49 notes per minute at the normal unstructured level, but only 31 notes per minute at the difficult unstructured level (see Table 11.3). This would seem to indicate that skilled participants are only able to maximise their performance tempo advantage over less skilled participants in the context of material that is within the bounds of their normal experience i.e. their superior performance with normal unstructured material does not confer further advantage with more obscure and demanding material.

11.3.3 Summary and conclusion

Despite its limitations, this study is valuable in that it provides evidence that participants who are broadly equivalent at fully rehearsed instrumental performance may differ significantly in terms of sight-reading related motor activity. This finding questions the common assumption within previous cognitive research that the origins of sight-reading skill difference are principally perception related. However, whether the variation in the recorded performances are due to differing dependencies upon visual feedback, differing abilities at unrehearsed movement, or a combination of these factors, awaits further research. Taken together, the sub-skill studies indicate that both perceptual and motor ability are strongly related to sight-reading skill. There are a number of ways in which such a finding can be interpreted, but the group-based analyses carried out so far are not capable of assessing their merits empirically. The next chapter considers these interpretations, and undertakes a more detailed exploration of individual participant data to see what empirical support might be available for them there.

Chapter 12

Analysis of individual participant data for the sub-skill studies

12 Analysis of individual participant data for the sub-skill studies

12.1 General introduction

The group-based ANOVAs have indicated that, overall, the skilled group are significantly faster than less skilled group in terms of both perceptual and motor sub-skill. Whilst this is a valuable finding, it unfortunately only serves to define the differences between the groups, not to explain them. Perhaps the apparently strong association of sub-skills is evidence of a shared representation; for example, geographical and spatial representations would appear to be implicated at the interface of perceptual and motor processing. Or maybe another external factor constrains them both equally, for example the amount of sight-reading practice that has been undertaken, or ability at sub-skill coordination. It is also possible that the association might disguise a more causative individual role for one or other of them. For example, if a musician has a problem with one of these sub-skills, it will not only hamper further sight-reading skill development, but may also limit the development of the other sub-skill area, unless it can be improved within another performance context.

Although the grouped ANOVAs are not able to provide insight into the origins of sight-reading skill variation, it is possible that an investigation of individual participant sub-skill data may prove more enlightening. Some individual results may exhibit significant variation from the general trend, from which more specific understanding of factors and mechanisms determining sight-reading skill might be gained. Something that might be particularly helpful in this regard is the fact that there is a distinct overlap in the group scores on both sub-skill studies. For example, if a less skilled participant only scores at a skilled level on the motor skill study, it might suggest that the reason for their slower reading ability is perception-related. It is also possible that there may be interesting patterns within the data of each skill group that may provide information about why skilled and less skilled readers differ from others within their skill group. Overall, therefore, an analysis of individual participant data similar to that undertaken in Chapter 9 would appear to be a potentially valuable exercise. The speculative nature of the approach needs to be emphasised once again. Final conclusions clearly cannot be drawn, bearing in mind that the samples are only small, and that the validity and reliability of the individual data cannot always be assured. However, the value of the investigation is, as before, seen primarily in terms of stimulating ideas towards which further research can be directed

This chapter is divided into two sections. The first is taken up with the individual participant analysis that has just been proposed. The second is concerned with exploring participants' introspections concerning the mental processes and representations used in carrying out the two sub-skill study tasks.

12.2 Sub-skill analysis

12.2.1 Introduction

This analysis involves correlation-based and descriptive examinations of the motor and perceptual sub-skill task data, considered in the context of the monophonic preview study results. As with the equivalent analysis in Chapter 9, only asymptotic performance has been considered because it allows for a greater combining of data across levels, thereby maximising reliability. For each individual participant, the perceptual sub-skill study is represented by a single, robust mean value combining their results (mean response time per note) at the 3, 4 and 5-note sequence levels together with both structure levels. The figure used for the motor sub-skill study is the participant's mean tempo for the entire experiment. The data for the latter study are response rates (larger values representing greater skill) whereas those for the former are response times (smaller values representing greater skill), and so to facilitate a more intuitive correlation analysis, the motor sub-skill data have been transformed into inter onset intervals. The monophonic preview data remain the same as employed in Chapter 9 except that they too have been transformed into inter onset intervals for the same reason.

Participants retain the same identity labels as before, S for skilled and L for less skilled, however the rank order on the monophonic preview study obviously now relates to ascending mean asymptotic inter onset interval size, within skill group. Participant L3 did not carry out the motor sub-skill study, and she has therefore been excluded entirely from this analysis. As in the previous individual participant investigation, the only analytical statistical measure that has been used is the

Spearman rank correlation coefficient, with a significance level of 5%. This correlation method has again been chosen because it is more appropriate given the different ranges of the distributions and the presence of outlying values.

First of all, the individual data of less skilled participants will be explored, followed by those of skilled participants. Finally, both groups' data will be considered as a whole, particularly with a view to gaining further insight into why the less skilled readers are slower than the skilled readers at monophonic sight-reading.

12.2.2 Less skilled participants

The data for less skilled readers from the two sub-skill experiments are presented in Table 12.1 and Figure 12.1, together with those of the monophonic preview study.

Table 12.1

Response times (ms) for individual less skilled participants representing mean asymptotic performance levels on monophonic preview, motor sub-skill and perceptual sub-skill experiments (rank position in parentheses)

Participant	Motor Sub-skill	Perceptual Sub-skill	Monophonic Preview
L1	855 (7)	548 (1)	214
L2	633 (3)	609 (2)	248
L4	840 (6)	595 (3)	284
L5	693 (5)	723 (5)	288
L6	519 (1)	654 (4)	288
L7	939 (8)	696 (6)	300
L8	1203 (10)	800 (7)	302
L9	1101 (9)	893 (10)	319
L10	623 (2)	863 (9)	388
L11	648 (4)	800 (8)	392

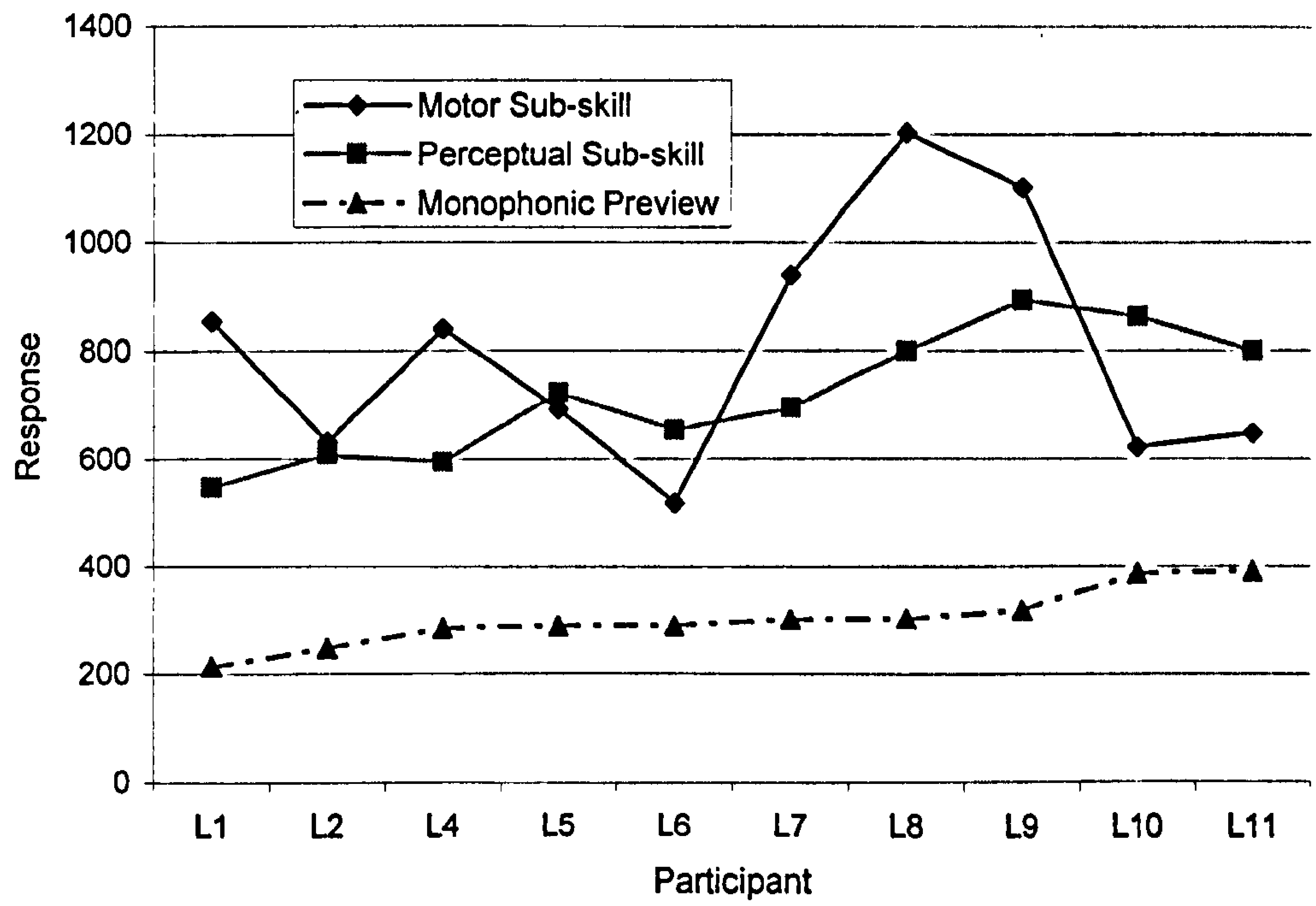


Figure 12.1
Response times (ms) for individual less skilled participants representing mean asymptotic performance levels on monophonic preview, motor sub-skill and perceptual sub-skill experiments

In Figure 12.1 the data points for each study have been joined to make them easier to distinguish. The results are evenly distributed throughout their ranges, with no evidence of any sub-grouping. A very strong correlation is found ($r_s = 0.90$, $p = 0.0004$) between the perceptual sub-skill and monophonic preview study data. This contrasts with the absence of any correlation between the motor sub-skill and monophonic preview results ($r_s = 0.04$, $p = 0.90$), and the two sets of sub-skill results themselves ($r_s = 0.15$, $p = 0.68$). In interpreting these correlations, it is important to keep in mind what the motor sub-skill data represent: a relative measure of visually unmonitored, unrehearsed motor ability. The measure is relative in nature because, as discussed previously, it would seem likely that the responses of participants were typically not maximised; their responses have been assumed, though, to be generally in proportion to their maximised response capabilities. With the measure interpreted in this manner, the fact that all participants performed more slowly on this study than on the monophonic preview study, which defines a minimum level for visually aided unrehearsed motor skill, is indicative of the constraint that the prevention of visual monitoring placed upon their output. However, the extent to which unrehearsed motor skill and dependency upon visual monitoring are individually responsible for the rank ordering in motor sub-skill study performance cannot be determined from the evidence available. Therefore, although the lack of correlation between motor sub-skill and complete task performance suggests that for less skilled participants, unrehearsed/visually unmonitored motor ability is not important in determining the level of monophonic sight-reading skill, it is possible that a relationship may exist for one of the factors individually. To enable further insight into this, the motor sub-skill study needs to be repeated with a condition allowing visual monitoring of output.

The lack of a visually monitored experimental condition also makes it difficult to interpret the strong correlation between the perceptual sub-skill and the monophonic preview study results. If neither skill at unrehearsed movement nor dependency on visual monitoring are associated with the level of complete task performance, the correlation would point to perception-related ability being the principal constraining factor on monophonic sight-reading, participants generally having sufficient, or indeed surplus, output skill with which to express their input-related capabilities. On the other hand, if there is an association between one of the individual motor factors and the complete task performance, clearly perceptual and motor sub-skills would then both be implicated in the patterning of sight-reading performance within the less skilled group. In such a case, it is possible that, as discussed earlier, both sub-skills may simply be reflective of sight-reading ability, and not fundamentally determining of it.

12.2.3 Skilled participants

The data for skilled readers from the two sub-skill experiments is presented in Table 12.2 and Figure 12.2 together with monophonic preview study results. The data of the perceptual sub-skill experiment fall into two main sub-groups, except for one participant. Six out of the ten participants (S1, S2, S3, S5, S7, S8) form the faster group, scoring between 453ms and 512 ms; three (S4, S6, S10) score between 668ms and 705ms, forming a slower performing group; the exception is a single individual who scores at a medium level 565ms (S9). In contrast to the less skilled participants, there is no evidence of a correlation between these perceptual sub-skill data and those of the monophonic preview experiment ($r_s = 0.20$, $p = 0.58$). It would appear from the

Table 12.2

Response times (ms) for skilled participants representing mean asymptotic performance levels on monophonic preview, motor sub-skill and perceptual sub-skill experiments (rank position in parentheses)

Participant	Motor Sub-skill	Perceptual Sub-skill	Monophonic Preview
S1	541 (6)	465 (2)	160
S2	412 (2)	512 (6)	178
S3	537 (5)	491 (5)	180
S4	367 (1)	705 (10)	187
S5	448 (3)	478 (4)	194
S6	619 (9)	668 (8)	220
S7	502 (4)	474 (3)	222
S8	637 (10)	453 (1)	224
S9	544 (7)	565 (7)	235
S10	564 (8)	679 (9)	243

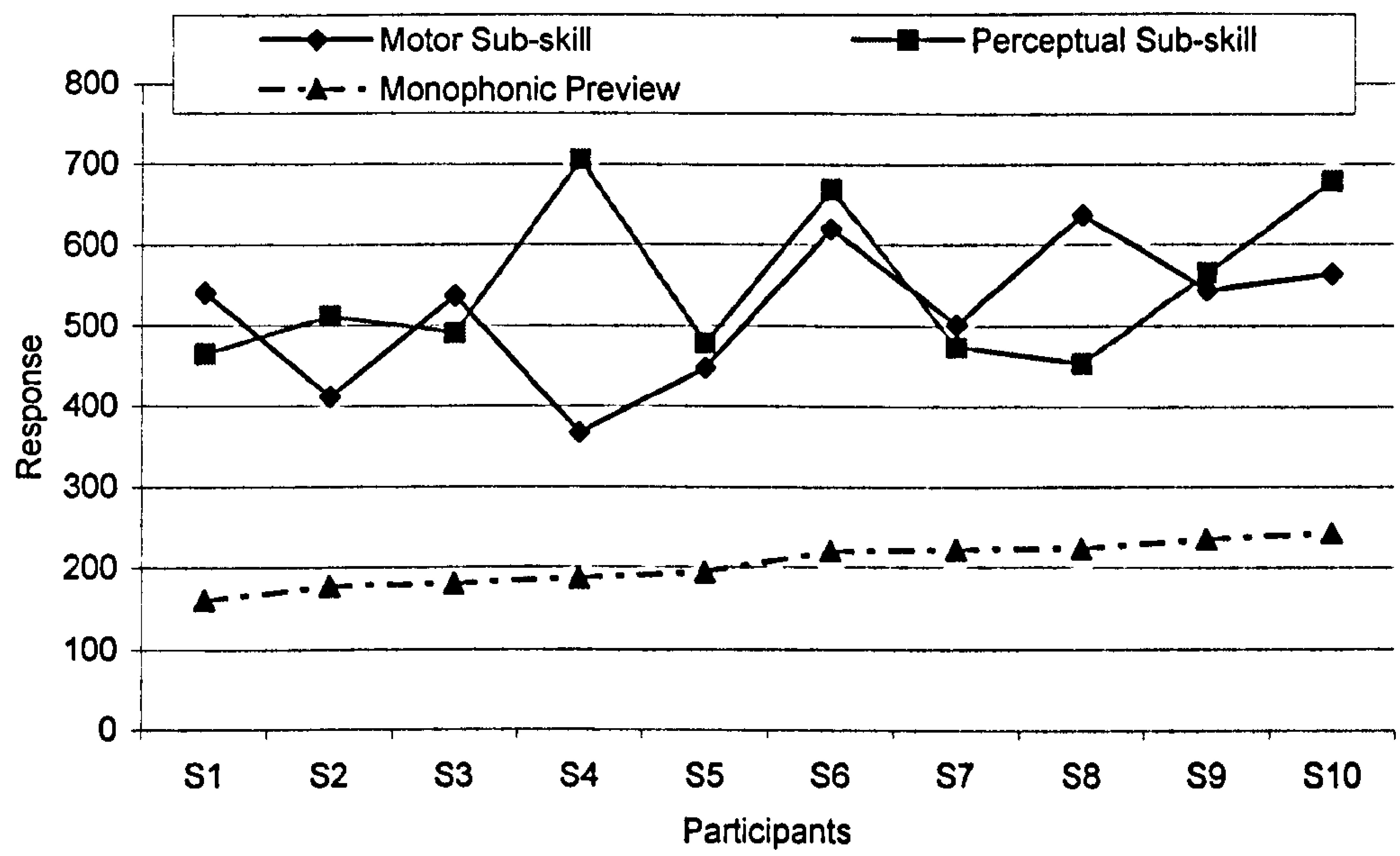


Figure 12.2

Response times (ms) for skilled participants representing mean asymptotic performance levels on monophonic preview, motor sub-skill and perceptual sub-skill experiments

results that fast speeds on the sub-skill study are neither necessary nor sufficient for high levels of expert monophonic sight-performance. For example, S4, one of the fastest participants on the monophonic preview study, is the slowest on the perceptual sub-skill task; S8, on the other hand, is the fastest performer on the latter task, and yet scores towards the bottom end of the skilled group with the former. This contrast between the less skilled and skilled data can perhaps be accounted for by proposing that skilled reading depends upon attaining a certain level of perceptual sub-skill, with ability beyond this less useful or important than other factors in determining actual sight-reading skill level. Such factors may include, for example, motor skill, performance monitoring and the coordination of input and output processes. Some limited evidence for this is provided by the fact that the correlation between the motor sub-skill data and the monophonic preview data approaches significance ($r_s = 0.56$, $p = 0.09$). However, no insight is available into the relative importance of skill at unrehearsed movement and dependency upon visual monitoring in determining this association.

12.2.4 Analysis of the less skilled data considered in the context of the skilled data

For both sub-skill studies, the results of L8 and L9 fall well outside the skilled group range, something that is entirely consistent with the findings of the sub-skill study ANOVAs. Their results therefore cannot provide any insight into what is acting to constrain their sight-reading performance level. The remainder of the less skilled group, however, all have at least one sub-skill result within or very close to the skilled group range, and a consideration of these results might enable more specific explanations of less skilled sight-reading performance to be determined. To facilitate this analysis, the skilled and less skilled data from Figure 12.1 and Figure 12.2 are presented together in Figure 12.3. On the perceptual sub-skill experiment, no less skilled reader performed at levels within the range of the fastest group of six skilled readers (defined earlier), but five less skilled readers scored within the range of the skilled group as a whole: L1, L2, L4, L6, and L7 (with L5, just outside of it). On the motor sub-skill experiment only three less skilled readers performed within the range of skilled readers: L2, L6 and L10 (with L11 just outside). Only two less skilled participants scored within the skilled range on both sub-skill experiments – L2 and L6. L2 performed well within the skilled range on the perceptual sub-skill study and just within it for the motor sub-skill study, results that are consistent with this participant's borderline skilled performance level on the monophonic preview experiment. L6 also scored within the skilled range on both sub-skill experiments, gaining the eighth fastest score for all participants on the perceptual sub-skill study and the seventh

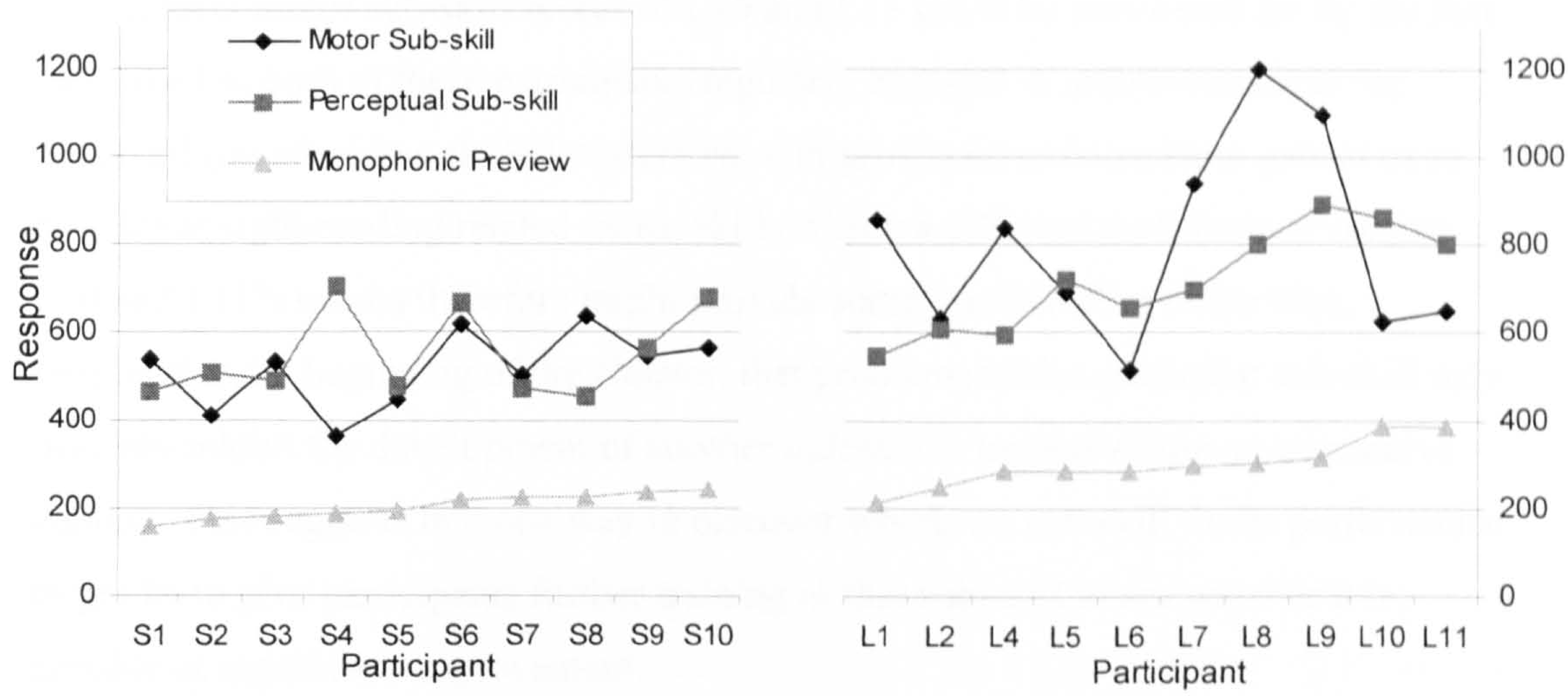


Figure 12.3
Response times (ms) for all participants representing mean asymptotic performance levels on monophonic preview, motor sub-skill and perceptual sub-skill experiments

fastest score on the motor sub-skill study. With these results, it might be considered that L6 had the potential for a skilled level of complete task performance, and yet he only managed an average less skilled tempo on the monophonic preview study. It is therefore possible that another factor was acting to constrain his performance here; for example, he may have had difficulties with performance monitoring, input/output coordination, or perhaps was experiencing interference between the two sub-skill area mechanisms. Another possibility could be that L6's skilled level of output on the motor sub-skill experiment related to particular ability at non-visually monitored performance, and that he was less able at the unrehearsed element of the task. It would seem unlikely that this is the case, though, because he was a regular improviser at the keyboard. It would seem worth mentioning, also, that this participant had his own theory about his slower sight-reading. He claimed to gain strong auditory representations from the score, which he felt interfered with, rather than informed, his performance.

Like L2, L10 and L11 scored on the border of the skilled range on the motor sub-skill study, however aside from L9, they were the slowest of all participants on the perceptual sub-skill study. This evidence may indicate that their output on the monophonic preview study was constrained by their weaker perceptual skill. The skilled level motor sub-skill scores of L10 and L11 could be accounted for by the fact that, like L6, both of these participants regularly engaged in improvisation at the keyboard (no other less skilled readers were improvisers) and therefore gained extra practice at sight-reading related motor skills within a different performance context. L10 and L11's results therefore might provide some confirmation of the idea, proposed at the beginning of the chapter, that problems with a particular sub-skill may possibly inhibit the development of another unless it is trained within an alternative context. This suggests that one way to discover whether a sub-skill limits performance might be to give participants further training in that sub-skill to see whether it is capable of significant improvement.

L1, L4, L5 and L7 are either within or on the border of the skilled range on the perceptual sub-skill experiment, but considerably outside of this on the motor sub-skill study. This may be an indication that factors relating to their motor sub-skills particularly hinder the sight-reading ability of these four participants. The result of L1, however, would seem to be anomalous considering her clearly skilled level of monophonic complete task performance. It could however be that her inferior motor sub-skill study output points to a particular need for the visual monitoring of movement and not to less effective unrehearsed motor activity. This participant's performance on the monophonic preview study may therefore have depended on a greater looking down at the hands than members of the skilled group who played at a similar tempo. Some confirmation of this may be provided by the decline in her performance with more complex material relative to the other members of the less skilled group, discussed in Chapter 9. It is possible that the lighter task demand of monophonic performance afforded considerable opportunity for visual monitoring to take place, but that this could not be sustained in the context of more complex output. If this is the case though, it would seem to contradict the experience of L8, who despite being the slowest performer on the motor skill study, improved her relative sight-reading performance level with more complex material to reach nearly the skilled group range on the four-part study (again, see Chapter 9). It is clearly not possible to account for the apparently contradictory performance of these two musicians from the current data.

12.2.5 Summary and conclusion

As has just been indicated, this analysis has raised more questions than answers. This was perhaps inevitable given the inability to discern the relative importance of unrehearsed motor ability and dependency on visual monitoring in determining the pattern of variation in the motor-sub skill data, something that is clearly a priority for future research to consider. Although the findings have typically been speculative because of the limited nature of the data, a sufficient amount of potentially meaningful individual variation would seem to have been demonstrated within the results to warrant a continued research focus upon the abilities of individuals, and not simply skill groups. The less skilled data shows evidence of participants falling into a number of sub-categories, each apparently requiring a different explanation for skill level. It would appear that for some participants the limiting factor is either perceptual or motor-related, but that for others, factors as yet undetermined, perhaps related to the coordination of sub-skills, may hinder overall skill development. So again, as was seen in the equivalent analysis in Chapter 9, the evidence suggests that sight-reading ability is dependent upon multiple, variable component mechanisms, and hence requiring of a rather more complex explanation than previous research and theoretical understanding has tended to consider. If this is the case, the priority for new research needs to be to gain a more detailed understanding of the 'anatomy' of the skill and its sub-skills, and of the typical variation in that 'anatomy'. Only in the context of such knowledge can a truly viable theoretical understanding begin to be developed. It is clearly essential that such research encompasses a broader ranger of notational complexities than has been attempted here, to gain a more complete picture of the sub-skill differences of skilled and less skilled readers and how these influence their complete task performance.

12.3 Introspections of mental processes during the sub-skill studies

12.3.1 Introduction

This analysis seeks to gain further insight into how musicians vary in their performance of sight-reading through a consideration of the conscious mental representations that they felt that they had used to encode materials during the performance of the sub-skill studies. There is much debate among psychologists about the origin, nature and functions of such mental imagery, and how it relates to unconscious mental processing. However, such issues do not need to be addressed in detail by the current analysis because the focus is not so much on understanding imagery itself, but rather using it as a window through which to further explore basic sub-skill variation amongst the participants.

There would appear to have only been a limited amount of published research into mental imagery and sight-reading aside from the work into auditory representation discussed in the literature review. In an early piece of work, Bean (1938) studied musicians' visual, auditory and kinaesthetic memory through a series of musical tasks and also asked participants to identify the type of mental imagery (again visual, auditory and kinaesthetic) that they typically depended upon. Whilst some participants seemed to be aware of using all three types of imagery equally, for others one type predominated. Furthermore, it was sometimes the case that a dominant representation coincided with greater skill at the analogous task. It clearly does not follow from this that it was necessarily the use of the imagery that was causative - the production of accurate images within a particular modality presupposes the possession of equivalent real-world skill. Other research has also indicated that an inability to imagine does not necessarily imply a lack of such skill. In Wolf's study of expert sight-readers (Wolf, 1976), already considered in the literature review, three of his four readers considered themselves to be dependent upon a single dominant imagery form, one tending to visualise the keys to be played, one using auditory imaging to plan movements, the other having an apparently precise kinaesthetic image of the keys to be performed. Although the reader who depended upon auditory imagery did not appear to use kinaesthetic imagery, as an expert, there can be no doubting his actual kinaesthetic skill. It needs to be considered why his imagery was lacking with regard to this particular modality.

Work in the area of visual imagery has suggested that approximately 10 to 12% of the population are unable to form conscious visual representations (Abelson, 1979). It is therefore possible that individual difference may similarly be responsible for the variation in imaging modalities displayed by skilled sight-readers. However, it may also be that such behaviour is simply learned. More recent musical consideration of mental imagery has been in the context of the study of mental performance rehearsal (for example, Connolly & Williamon, (2004)), where it would appear that the learned nature of conscious representations has generally been emphasised, although it is not denied that the modalities may resonate with different individuals to varying degrees. Of course, it remains possible that failure to produce images within a particular modality might be due to a lack of real-world ability. For example, Wolf's reader who depended upon kinaesthetic imagery may have had weaker aural skill, but compensated for any deficiency by using kinaesthetic ability to a greater extent. Overall, therefore, it would seem that underlying skill is necessary but not sufficient for the production of an analogous conscious mental representation.

No previous published work appears to have investigated whether skilled and less skilled sight-readers differ in their experience of task-related imagery. As has previously been explained, for this research participants were engaged in a short conversation after each of the sub-skill studies and asked to describe the principal representational strategies they considered they had used in their performances. Because there is no way of determining the accuracy of participants' recollections, the validity of introspective data is always going to be open to question. Furthermore, an issue relating to this particular research is that the individual reflections are perhaps not as considered and detailed as they might be. Participants had already committed considerable time to the performance aspects of the research and making further significant demands upon them in this area was seen as inappropriate. Nonetheless, they were generally very clear about the strategies they had employed. Whilst these concerns are not sufficient to prohibit analysis, it obviously needs to proceed with due caution. It also needs to be borne in mind that in view of the limited understanding of mental imaging in general, not simply within sight-reading, the results may not be capable of precise interpretation.

This analysis begins with a consideration of the memorisation and preparation strategies used for the motor sub-skill study. Before undertaking this it would be beneficial to provide some information relating to the aural abilities of the participants. Participants were required to categorise themselves as possessing either perfect pitch, excellent, good or poor aural skill. Such self-assessment is not ideal, but there would seem no substantial reason to doubt these experienced musicians' assessments of their own abilities. The participants fell into three categories:

Perfect pitch	S1, S2, S4, S6, S9, L1*, L5
Excellent aural skills	S3**, S5, S7, L6, L11
Good aural skills	S8, S10, L2, L4, L7, L8, L9, L10

* L1's perfect pitch was acquired and apparently not completely fluent

** S3 claimed that she had very near to perfect pitch

Some participants claimed to produce clear auditory representations as they input the materials from the experiments. These were all the skilled readers who had perfect pitch (S1, S2, S4, S6, S9) and also S3, L6 and L11 who possessed excellent aural skills. No readers from the 'good aural skills' group claimed to be conscious of any audiation.

12.3.2 Motor sub-skill study

There were three principal strategies that participants appeared to use to encode the key and finger number stimulus into short-term memory and prepare their performances. These relate to the dominant representations that participants felt they were using, and so do not necessarily preclude the use of others.

1. Strategies dominated by kinaesthetic imagery combined with explicit audiation of the stimulus. (S1, S3, S4, S6, L11). Participants using this strategy possessed either perfect-pitch or excellent aural skills. The separate note position and finger number components of the stimulus appear to have been combined into a single embodied stream, either a rehearsed hand position or series of hand movements. (My use of the term 'kinaesthetic imagery' encompasses both position and movement-based representation).

2. Strategies dominated by kinaesthetic without explicit audiation of the stimulus. (S5, S7, S8, S10, L2, L10). Participants using this strategy were those with excellent and good aural skills. Although these individuals were not aware of forming any detailed auditory representations, all but one of them claimed to have detected their performance errors in the study by aural means. Only L2 claimed not to have been aware of any aural feedback at all, this participant being insistent that her error detection was entirely kinaesthetic-based.

3. Strategies based upon auditory imaging of notes and lists of numbers for fingers. This might appear an unlikely strategy, but two skilled (S2 and S9) and one less skilled participant (L6) nonetheless claimed to have used it consistently, all having either perfect pitch or excellent aural skills. All three claimed categorically to have had no sense of any mental rehearsal based upon kinaesthetic imagery. The note and finger number stimulus components would therefore seem to have been stored as two separate representational streams prior to performance.

The remainder of the participants (L1, L4, L5, L7, L8 and L9) used a range of non-specific, *ad hoc* strategies lacking in any consistent representational focus, involving a search for visual and musical patterns, numbering patterns and the use of mnemonics. These six participants were all from the less skilled group, and apart from L1, had good aural skills. None of these participants were aware of the use of any kinaesthetic imagery. Neither were they conscious of any explicit auditory representation, although as with strategy number two above, they all considered that they had used aural means to detect performance errors. It would appear that for these participants, notes and finger number storage involved separate representations, but without the clear representational focus of S2, S9 and L6 described in the previous paragraph.

The analysis provides confirmation of previous sight-reading research in three ways:

1. It affirms that skilled sight-readers do vary in terms of the principal imagery that they employ during sight-reading related activity.
2. Further evidence is provided that the experience of imagery within a particular modality presupposes analogous real-world skill. The use of kinaesthetic imagery by eight members of the skilled group is clearly consistent with their skill level. L2, L10 and L11's experience of kinaesthetic imagery is also consistent with their skilled output on the motor sub-skill study (L11 was just outside the skilled range). As far as auditory imagery is concerned, it was only participants who had perfect pitch or excellent aural skills who claimed to experience this, including one less skilled reader (L6). All the skilled readers employed one of the three principal strategies listed above, as did the four less skilled readers just mentioned. The remainder of the less skilled group, on the other hand, appear to have had recourse to neither kinaesthetic nor auditory imagery, and seem generally to have chosen their approach on a test-by-test

basis. The lack of kinaesthetic imagery amongst these six less skilled participants is consistent with the fact that they were the slowest performers on the motor sub-skill study. Their lack of auditory imagery is also consistent with their claim to have only 'good' aural skills. Whilst S8 and S10 demonstrate that this category of aural skill is quite sufficient for skilled reading, it would appear, at least from the current results, that it is not able to support the highest levels of skilled performance. For example, all the professional accompanists (S1, S2, S3 and S4) had perfect pitch or excellent aural ability. Therefore, the lack of strong aural ability of the six less skilled readers, just discussed, may also have been a contributing factor to their slower sight-reading performance.

3. The analysis confirms previous research indicating that conscious representations in a particular modality do not necessarily accompany the equivalent real-world skill. For example, S2, S9 and L6 although obviously possessing strong kinaesthetic ability (L6 was the fastest less skilled participant on the motor sub-skill study) were not aware of using kinaesthetic imagery. It appears that not only did they not even consider using such imagery during the study, it was, surprisingly, a type of representation that none of them were familiar with in the context of normal sight-reading. It is however possible that these participants were capable of employing such imagery but merely did not think to do so given the particular nature of the stimulus. Maybe the materials used were impoverished in pianistic terms, or perhaps a film of actual movements to copy might have been more successful in stimulating its use. With regard to auditory imagery, the possession of perfect pitch or excellent aural ability was no guarantee of participants' ability to experience this. For example, S5, S7, L1 and L5 all fell into these two categories but claimed not to have not been aware of audiation during the experiment.

Perhaps the most significant new finding that these results have provided is the lack of any focused use of imagery amongst the majority of the less skilled group. The inability of these six participants to form either kinaesthetic or auditory imagery together with the *ad hoc* nature of their encoding strategies suggests that they lack a secure and efficient means of organising memory storage in relation to sight-reading performance. Their reported mental strategies are consistent with the inferior performance of less skilled readers compared to skilled readers across the research as a whole i.e. their slower processing, their typically larger number of errors and the possible evidence of smaller task-related short-term memory capacity. Whilst it is possible, as discussed above, that the inability of the six less skilled readers to make use of kinaesthetic or auditory imagery may be reflective of their level of sight-reading skill, it could also be, as Abelson (1979) has indicated in relation to visual imagery, that factors to do with their individual make-up may be preventing them from forming these types of conscious representation. If this is the case, it may be that their inability to image in one of these modalities may actually be constraining their skill development. The fact that each member of the skilled group made use of at least one of these types of imagery may indicate their use to be essential in achieving an organisation of memory mechanisms capable of supporting skilled reading.

Previous research that was discussed earlier mentioned that that some sight-readers made particular use of visual imagery, but participants in this study made little mention of representation within this modality. One reason for this might be that they were less aware of visual imagery because the stimulus itself was presented as a visual representation of the piano keyboard. It is also possible that visual imagery was implicit within the formation of the other representations. Clearly, though, the data concerning conscious representation provided by my participants is too rudimentary for any firm conclusions to be drawn about this, or indeed any of the issues that have been considered in this section. However, the findings would seem to be sufficiently consistent with other results within the research as a whole to warrant a more sophisticated investigation.

12.3.3 Memorisation strategies used for the perceptual sub-task experiment

There was less variation in strategy between the participants and skill groups in relation to this study. All who had described strong auditory representations for memorisation on the motor sub-skill study continued to use this as their primary strategy, except for S3, who felt she was mainly carrying out the trials using visual imagery, and identifying marker notes and patterns of intervals around them. Only two participants (S6 and S8) claimed to have used any embodied imagery in their performance. All other participants appear to have employed strategies equivalent to that of S3, with some participants from both groups explicitly ascribing letter names to some of the keys. The fact that so little reference was made to embodied imagery may merely indicate that there was no requirement for perception and cognition to drive actual motor activity. However, it is also possible that the memory demand was not great enough to invoke such representations, or that the materials were not sufficiently pianistic. The fact that visual imagery appears relevant to performance in this study but not to the motor sub-skill experiment would seem to confirm the explanation given earlier that the lack of a transcriptional element either made visual imagery less necessary, or else made participants less aware of using it. Another possibility is that the memory demands of that study may have been too great for visual imagery alone to suffice, therefore requiring the more dominant use of other modalities.

Chapter 13

Conclusion

13 Conclusion

13.1 Introduction

This conclusion is divided into three main sections. The first, and principal, section gathers together and reviews the main findings of the different studies, considering the extent to which they have added to an understanding of sight-reading and discussing the theoretical implications of this. Some directions for future research are also suggested. The second section discusses the experimental methodologies, examining their effectiveness and possible ways in which they might be improved for future study. The final section explores whether the research in this thesis can offer any pedagogical insights into how sight-reading skill might be improved.

To set this closing discussion in context, it would be useful briefly to revisit the original reasons for undertaking this research. As was made clear in the introduction and the literature review, there were two principal motivating factors. The first was the perceived need for research to provide more detailed knowledge into foundational aspects of the task. The second was concern over what I have termed the patterning account, the dominant account within music psychology of why skilled musicians differ in sight-reading ability. Proponents of the account claimed the research evidence to point primarily to a structure-based explanation, skilled readers being able to 'circumvent limitations of the human information-processing system' (Lehmann *et al.*, 2007, p. 114) by processing larger groups of notes more quickly than less skilled readers, courtesy of their greater sensitivity to musical patterning in the score. However, my own reading of the literature raised significant doubts about the extent to which the associated research could confidently be considered to support these conclusions. For example, the signs were that skilled readers did not necessarily process notes in larger groups than less skilled readers, and even when they did, there appeared to be no reason to conclude that the cause was structural in nature, rather than perceptual or motor. Furthermore, the little research available on the issue actually gave no indication that less skilled readers were any less sensitive to musical structure than skilled readers. In my view an explanation of skill difference based upon more elemental perceptual or motor factors appeared more in keeping with the evidence, an explanation that I have termed the perceptuo-motor account. The research in this thesis was therefore set the dual aim of gaining greater insight into fundamental aspects of skill and attempting to clarify the situation with regard to the patterning and perceptuo-motor accounts.

As we have seen, the subsequent experimental work has failed to provide support for the claims of the patterning account. Clearly, any research that questions a long-standing and widely accepted theoretical account needs to be viewed with an element of caution, and to undergo detailed examination to uncover possible methodological weaknesses and faulty interpretations. However, in reality, my studies would appear to have unearthed little that is actually new or controversial; overall, they have simply served to confirm and elucidate the empirically based concerns that originally inspired them. To emphasise this, as I draw together the main strands of my own research, I will document, where appropriate, the previous findings for which they provide confirmation.

13.2 Main research findings

13.2.1 Introduction

My experimental work was an attempt to provide answers to a number of basic questions that would enable an assessment of the relative merits of the patterning and perceptuo-motor based accounts to be made. These questions are presented below and provide a useful focus and order for this final summing up of research findings.

1. Do skilled readers use greater preview and larger perceptual units than less skilled readers and if so is this the source of their superior performance?
2. To what extent is the performance level of skilled and less skilled readers dependent upon structural perception?
3. To what extent are the effective preview and perceptual unit sizes of skilled and less skilled readers dependent upon structural sensitivity? Are these measures constrained by short-term memory capacity, or are other factors responsible for this?
4. Do skilled and less skilled readers differ significantly with regard to basic perceptual and motor sub-skill?
5. Are sight-reading related component skills strongly correlated amongst musicians suggesting an overall simple account of skill difference? Or is there

sufficient variation to indicate that there may be a more complex range of causes?

To provide a meaningful overview and assessment of the research findings, this section involves repetition of earlier analysis and discussion, but the aim has been to set these now more firmly within the context of the work as a whole.

13.2.2 Size of preview and perceptual unit

The empirical results of this thesis have consistently demonstrated that the skilled participants did not, as proposed by the patterning account, depend upon greater preview for their skill superiority over the less skilled participants. With regard to the controlled preview studies this is most clearly seen in the results for two-part structured material, where the two groups obtained identical maximum effective preview spans (see Table 13.1, which provides span data for all three studies).

Table 13.1

Mean maximum effective preview span, in beats, for each skill group with monophonic, two-part and four-part material (mean associated tempo in parentheses, in beats per minute). Two-part and four-part data represent individual structure levels; monophonic data are combined across structure levels.

		Monophonic	Two-part	Four-part
Skilled group	Structured	6 (299)	4 (193)	3 (122)
	Unstructured		4 (179)	3 (108)
Less skilled group	Structured	4 (204)	4 (106)	2 (67)
	Unstructured		3 (92)	2 (56)

The skilled group therefore achieved a faster maximum speed than the less skilled group by being more efficient at processing the same quantity of preview. Although the skilled group did make use of greater preview than the less skilled group with all other categories of notation, it was always the more efficient processing of levels of preview shared with the less skilled group that defined their skill superiority; the extra preview was only responsible for relatively limited, additional performance gains. This finding is consistent with the results of earlier research by Sloboda (1974, 1977) and Gilman and Underwood (2003). The fact that Sloboda's skilled readers recorded smaller eye-performance spans with less structured material than with structured material without necessarily a concomitant decline in performance level, suggests that his skilled readers' typical span advantage over less skilled readers on the latter type of material (skilled readers = 7 notes, less-skilled readers = 4 notes) was not as essential to reading skill difference as he had considered. Gillman and Underwood quantified the effective preview of their skilled and less skilled groups as 4-5 beats and 3¾ - 4¾ beats respectively with 3-part material, figures very similar to my two-part data.

What appears to be a completely new finding is that the skilled group actually needed a smaller amount of preview than the less skilled group to achieve a faster level of performance. For example, using only three beats of preview on the monophonic study, they were able to play approximately 10% more quickly than the less-skilled group's mean asymptotic tempo. Such a small level of effective preview might be considered to contradict Sloboda's considerably larger skilled eye-performance span measure recorded in the context of skilled and less skilled readers sight-reading at the same tempo (discussed in the last paragraph). However, my measure would seem to represent the level of preview necessary for basic note finding, whereas his may incorporate the additional preview needed for fluent musical performance, for example the planning of fingering and performance expression. The possibility that the size of effective preview may vary for different task components emphasises the importance of researchers having a clear understanding of the meaning of their measures prior to data interpretation. Considered in the light of this analysis, it would seem possible that the proponents of the patterning account have overestimated the extent to which the skilled eye-performance span is essential to the note identification process.

In relation to the pattern-matching study, there is no evidence that the faster performance of the skilled readers depended upon a larger notational context; they consistently outperformed the less skilled readers at all sequence lengths, with both skill groups reaching their greatest efficiency of processing at the same sequence length. This evidence is in line with the results of the pattern-matching studies of Waters *et al.* (1997, 1998a), which found no significant difference in skilled and less skilled readers' patterns of performance, except for speed of response. Waters *et al.* considered that their research might indicate that skilled readers processed notes in larger groups than less skilled readers, however this interpretation would seem to attach too much importance to results that show only trends towards significance. Both skilled and less skilled readers in my pattern-matching study reached their most efficient level of processing at 3 or 4-note sequence lengths. This may indicate that these figures represent the typical size of perceptual unit used with the longer sequences, and also perhaps in monophonic sight-reading. Truitt *et al.* (1997) may provide some supporting evidence for this, their experienced and less experienced musicians recording perceptual spans of 3 to 4 notes during monophonic sight-reading. However firm conclusions cannot be drawn about this because of methodological differences between the studies.

13.2.3 Musical structure and performance tempo

On all three controlled preview studies, both skill groups demonstrated similar tempo responses to the structural distinction. On the monophonic study there was no significant effect of structure. On the other studies, participants performed more quickly with structured material, although the magnitude of the effect was not particularly marked (see Table 13.1). The simplest and most plausible interpretation of these results is that the two skill groups were similarly sensitive to musical structure within the score, pointing to some other factor or factors being responsible for the significant effect of skill, for example, perceptual or motor ability. However, by considering the effects of skill and structure to be the result of a more complex set of influences, a fundamentally patterning-driven interpretation of the variation in the skill groups' performances is possible. I refer the reader to chapters 6 and 7 for a detailed discussion of this interpretation, but essentially it proposes that the skilled readers may have been able to perceive a considerable degree of structure in the unstructured sequences to organise their performances. Whilst theoretically possible, such an interpretation would seem unrealistic in practice, however. It requires that the skilled readers gained their perception of structure from a narrow notational context i.e. preview levels that they shared with less skilled readers, something that would seem highly implausible with unstructured material. Indeed, skilled participants performed unstructured material significantly faster than less skilled participants even at small preview sizes where very little unambiguous structural information is available.

The fact that both groups of participants performed approximately 10% faster with structured material than with unstructured material on the pattern-matching study suggests that there was the potential for a structural effect on the monophonic preview study, but that both groups were able to achieve their fastest level of output without the aid of the extra patterning available to them in the structured sequences. Although the effect of structure was similar in absolute terms for both skill groups on the two-part and four-part studies, when one takes into account the large difference in their performance tempi (see Table 13.1), it was actually the less skilled readers who made the largest relative gains from the greater availability of musical patterning. At asymptotic tempi on both these studies the skilled group continued to perform structured material approximately 10% faster than unstructured material; however the less skilled group made gains of 15% with two-part material and 20% with four-part material. This suggests that with increasing task demand, the less skilled readers depended more upon structure than the skilled readers to achieve a maximisation of their performance, a finding that is completely contrary to the expectations of the patterning account, but entirely consistent with a perceptuo-motor based understanding of skill (see discussion in Chapters 7 and 8).

Although the finding that my less skilled group's performances are at least as sensitive to structure as those of my skilled group contradicts the patterning account, it is in fact in line with the very small amount of research evidence available in this area. One of the pattern-matching studies of Waters *et al.* (1997) involved some rudimentary structural disruption of sequences and revealed that skilled and less skilled readers did not differ significantly in their response to this. Also, in Sloboda's eye-performance span research, the spans of less skilled readers expanded and contracted in response to structural cues in a manner proportionally equivalent to skilled readers (Lehmann *et al.*, 2007). Although there would appear to be no evidence to suggest that skilled readers are more sensitive to musical structure than less skilled readers, there are clear signs that experienced musicians are more so than novice musicians (Halpern & Bower, 1982; Waters et al., 1997), a finding that is consistent with work in the expertise paradigm that inspired Sloboda's early formulation of the patterning account (Simon and Chase, 1973). The lack of evidence supportive of the patterning account itself suggests, as argued in the literature review, that Sloboda's original use of Simon and Chase's paradigm to attempt to explain sight-reading skill variation amongst experienced musicians (Sloboda 1974, 1977) was inappropriate. Insufficient consideration had been given to how the context of the less skilled readers' considerable musical knowledge and ability might inform their sight-reading performance. In other words, although they were certainly less skilled at sight-reading, they could scarcely be defined as novices at the task. As has been noted previously in the thesis, such failure to distinguish adequately between the skills of novice and experienced musicians with regard to sight-reading related activity has continued to be a problem within the domain, and is clearly something that future research and theory development needs to address.

The minimal tempo effects in response to the structural distinction question the patterning account proponents' contention, mentioned at the beginning of the chapter, that the perception of structure enables skilled readers to circumvent the normal limitations of information processing. Performance with structured material in my studies would instead appear to be very much constrained by the information processing systems of skilled readers - a stretching of system limits perhaps might be indicated, but there is nothing to suggest circumvention of them. Of course, it must be borne in mind that the tempo effects of the structural distinction only represent relative, not absolute, responses to structure, so my data might underestimate its importance. However, if this is true for skilled readers, a similar case could also be argued for the less skilled readers. It is also possible that the lack of a rhythmic element or explicit musical style might have prevented the proposed circumvention from happening. However, in relation to this latter point, skilled participants did generally perform very quickly with structured material and it would seem unlikely that any significant increase in sight-reading speed beyond this would be attainable.

With the lack of music-based research, discussion about patterning in the sight-reading literature often makes references to findings from research into alphabetic text (Lehmann *et al.,* 2007) The wisdom of this is questioned however by a comparison of responses to the structural distinction in my research with that of Shaffer (1976), considered previously in the literature review. Shaffer found that skilled typists achieved approximately a fourfold increase in speed by using normal text as opposed to text in which letters within words were randomised. In my preview studies, as we have just seen, the largest average gain in speed from the structural distinction was tiny in comparison - approximately 10% for skilled participants. It is possible, therefore, that in proposing that skilled readers have the ability to circumvent the limitations of their information processing systems, the proponents of the patterning account may have considerably overestimated the extent to which the graphical nature of musical notation is typically susceptible to structure-based encoding efficiencies.

13.2.4 Structure, preview and factors limiting to performance

On the surface, the maximum effective preview span evidence suggests that the greater availability of structure does not have much of an effect upon preview. The influence of structure was only visible in the less skilled group's performance of two-part notation, where they achieved one extra beat of preview on the structured level (see Table 13.1). The lack of an overall effect would seem consistent with the only limited tempo responses of the skill groups to the structural distinction. However, no firm conclusions can be drawn about this because with multi-part material the extent to which notes within the final beat of preview have been processed cannot be discerned. In other words, a structural effect may indeed be present with regard to preview, but simply not detectable with the current methodology. Therefore, it is possible that the difference in span obtained for less skilled participants with two-part material is evidence of the beginning of a trend, rather than merely a one-off result. This issue needs further research.

It has been clearly documented in the literature review that the patterning account considers the limiting factor upon the sight-reading performance of less skilled sight-readers to be their inability to use musical structure to increase short-term memory storage and hence preview, and that this is the case even with basic monophonic notation. The fact that the maximum effective preview of my less skilled group ranges from 6 to 8 notes (3 to 4 beats) with two-part material but is only 4 notes with monophonic material suggests that memory storage capacity was not a limiting factor on their monophonic performance. Considering, however, that they again make use of 8 notes of preview (4 beats) with four-part material indicates that an asymptote of storage has been reached, consistent with short-term memory capacity starting to be a limiting factor with two-part material. However, it is also possible that some other factor is responsible for the asymptote and so further research is required. Comparing the maximum effective preview spans in the same way for the skilled group finds no clear evidence consistent with short-term memory capacity being a limiting factor in their case. The fact that they have larger effective preview spans than the less skilled group with four-part preview may indicate that they are capable of greater levels of raw short-term memory storage.

A number of points emerge from these results. Firstly, together with previous research by Gilman and Underwood (2003), mentioned earlier, showing that skilled and less skilled readers are capable of very similar levels of effective preview with dual stave material (skilled = 4 - 5 beats; less skilled = 3 ¾ - 4¾ beats), they suggest that Sloboda's monophonic eye-performance spans do not represent the memory storage capacity of his skilled and less skilled readers, as he interpreted them in the development of the patterning account (Sloboda 1978b), but rather their relative rates of encoding recorded within a limited time-window. To be sure that his measures represented the former, he should have required his participants to also sight-read at slower tempi, providing longer time-windows for perception and therefore allowing an asymptote for storage to be quantified. Further research needs to be carried out into the eye-performance span to clarify this issue.

Secondly, it is important for work to be undertaken to identify the factors limiting performance for different levels of notational complexity. Two obvious possibilities that need to be considered, aside from memory storage capacity, are the speed at which material can be encoded and also skill at unrehearsed motor output. Gaining insight into the extent to which musicians have preview beyond necessary levels might provide some insight into these. For example, the availability of non-necessary preview might indicate that motor output is constraining overall performance i.e. further perception is available but not being utilised.

Thirdly, my maximum effective preview spans would seem to represent unsustainable memory demands, especially in the case of unstructured, more complex materials. For example, Sloboda's eye-performance span research has indicated that skilled readers cannot typically store in short-term memory more than 7 notes of tonally coherent notation, and that storage capacity declines when tonal structure is disrupted (Sloboda, 1974, 1977). One factor that might have facilitated information about a larger number of notes being stored in memory during the performance of my preview studies is the fact that, unlike Sloboda's methodology, the score was always available to participants, enabling memory to be continually refreshed. However, findings from research into touch-typing and text reading (see Chapter 8) suggest that detailed memorisation of the entire span may not be required. Evidence from touch-typing research (Salthouse 1986), has demonstrated that material about to be performed (between one and two keystrokes) has typically already passed into output buffers and therefore makes only minimal demands on memory processing. Studies into text reading (Rayner, 1998) have shown that material at the forward extent of preview may only undergo partial processing but nonetheless still be useful in terms of priming subsequent activity. In the context of sight-reading this might mean that simply perceiving the general direction of required hand movements could be effective for performance. This evidence suggests, therefore, that it might only be notes in the central area of effective preview that require detailed identification and processing in memory. However, further research is clearly necessary to confirm the extent to which these mechanisms are relevant to music sight-reading.

A final point to mention concerns the issue of short-term memory storage capacity as a limiting factor to sight-reading ability. Results for preparation time in the motor sub-skill study suggest the possibility that the skill groups may not differ in terms of absolute storage capacity, but that as the quantity of material to be processed increases, less skilled readers simply become relatively less efficient at storage compared to skilled readers. It may be, therefore, that it is not absolute memory capacity that is limiting to less skilled readers' performance for more complex musical materials, as suggested by the four-part preview study, but rather the availability of memory that is 'fast' enough to support further increases in performance speed. More work into the memory abilities of skilled and less skilled sight-readers is needed, both specifically to investigate this issue, and to increase general knowledge about their respective skill make-ups.

13.2.5 Sub-skill studies involving monophonic material

As well as demonstrating that skilled readers are generally faster at perception related activity, the perceptual sub-skill study provides insights into performance with difference sequence lengths and structural content that have been discussed in previous sections. The error rate for the two skill groups is very low on this task, indicating that the slower performance of less skilled participants was not related to any conceptual problem with task performance. Whilst the motor sub-skill study appears successful in demonstrating that the skill groups differed in terms of unrehearsed non-visually monitored motor ability, their slow performance tempi at the task compared to the monophonic preview study means that experimental conditions of structure and sequence length are not able to offer any detailed insights into the real-world abilities of the skill groups. This issue will be discussed further in the section relating to methodology. The considerably greater number of errors for both groups in this experiment compared with the controlled preview study confirms previous findings that at least some visual feedback is essential for normal performance (Banton, 1995). Without additional evidence relating to more visually assisted output, however, there is no way of clearly discerning the extent to which the difference in performance of the two skill groups at the task (or indeed the within-group variation) was the result of differing levels of unrehearsed motor skill or varying dependency upon visual monitoring. This is an important area for future research to focus on. It would also be important for sub-skill work to explore dual-stave and rhythmic material to gain a more complete picture.

A final point to mention in relation to motor skill is that the analysis of participants' fingering that was mentioned in Chapter 4 has not been carried out for this thesis.

13.2.6 Analysis of individual participant data

Although such analysis runs the risk of reading meaning into variation that may not be skill-related but simply a random experimental effect, it was never intended to be authoritative, the principal aim being to mine the data to gather ideas for future research to explore. Study of the controlled preview data reveals that whilst some skilled and less skilled participants demonstrate a strong relationship between their performances across the three notational complexities, many show particular ability, or particular weakness, with one type. This confirms the concerns that were raised in the literature review about using only a single measure to quantify sight-reading skill. One might have expected that particular weakness at the task would be associated with more complex materials, but this was not necessarily the case. For example, one less skilled participant was one of the slowest performers with monophonic notation, but very nearly reached skilled group levels of output with four-part notation.

Although there was substantial variation in the performance of some individuals across the three types of notation, there was little between-group overlap in scores. Two less skilled participants recorded results within the skilled group range, but only with monophonic notation; they were not able to maintain their performance level with more complex materials. Overall this indicates, that a skilled level of performance across the three types of notation studied was necessary for participants to consider themselves skilled sight-readers. The fact that the results of these studies, taken together, successfully categorise participants in line with their self-labelling also indicates that overall pitch playing ability is a sufficient determinant of skill group categorisation i.e. no reference needs to be made to ability at rhythm reading.

Turning to the analysis of the sub-skill study data, the grouped analyses indicate that both perceptual and motor ability are strongly linked to skill at the complete task. This finding is merely definitional, though - it does not provide any reasons for the different sight-reading abilities of the two groups. However, a greater overlap of results for the skill groups on these two studies than for the controlled preview experiments, together with considerable variation amongst the less skilled group's results, makes possible some limited, and necessarily cautious, detection of factors that might be constraining the performance of the less skilled participants. In particular, there is evidence to suggest that perceptual and motor ability can act individually as limiting factors. Less skilled participants scoring at skilled levels on both studies may indicate a problem with the coordination of sub-skill activity, or interference between the sub-tasks. Regarding skilled participants, it would appear from the results that fast speeds on the sub-skill studies are neither necessary nor sufficient for high levels of expert monophonic sight-performance. This suggests that beyond a certain level of sub-skill ability other factors become more important in determining skill level at the complete task, for example, the ability to coordinate sub-skills effectively.

In preparing performances on the motor sub-skill study, all the members of the skilled group described employing auditory imagery, kinaesthetic imagery, or both. This confirms the finding of previous research (Bean, 1938; Wolf, 1976) that musicians vary in terms of the type of imagery that they use in sight-reading related activity.

The same strategies were also in evidence amongst the less-skilled group, though six out of the ten who undertook the task stated that they experienced neither of these forms of imagery, but instead resorted to *ad hoc*, non-embodied methods of sequence representation. I am not aware of any previous research into the conscious memory representations of less skilled sight-readers, and so this lack of embodied imagery use amongst the less skilled group appears to be a new finding. However, it is not clear what the evidence points to. Perhaps the lack of any focused experience of imagery in the particular modalities simply reflects these musicians' less developed underlying aural and unrehearsed motor skill. However, it is also possible that other factors may hinder the generation of such imagery. If this is so, considering that all the skilled participants made focused use of either motor or auditory imagery, it could be that the availability of at least one of these forms of imagery is a vital component to the memory organisation of skilled sight-reading performance. The evidence indicates that the study of imagery in relation to sight-reading warrants further attention. It would be particularly enlightening to relate these findings to musicians' experience of imagery in relation to rehearsed performance.

Analysis of individual participant data reveals a complex pattern of individual variation across a range of component skills - perceptual, motor and representational - providing further evidence that understanding the reasons for the different sight-reading abilities of skilled and less skilled readers requires a consideration of a much wider range of factors than has typically been the case within the domain. Although grouped research is clearly important, the degree of within-group variation found amongst my participants suggests that to gain a better understanding of the 'anatomy' of the task and its variation within the population, a greater emphasis upon the study of the skills of individual musicians is needed. For this kind of research it is especially important that the methods and technology used are capable of providing valid and reliable data, and so I will now turn to an evaluation of my experimental approaches and how they might be refined for future work.

13.3 Methodology and technology

The particular type of controlled preview methodology used in this thesis has not, to my knowledge, been previously employed in the study of sight-reading, although it is a common and apparently successful approach in touch-typing research. Prior to the research, there was some concern over ability of the approach to provide valid and reliable data. Firstly, it was considered that the continual updating of the screen might be off-putting to participants and so prevent a true assessment of their abilities. The data, however, suggest there to have been no major problems in this area. For example, tempi at unlimited preview, in which all the notation is shown and there is no updating of the screen, are similar to controlled preview asymptotes. Secondly, there was the issue of the novel use of tempo as a performance measure within sight-reading. Of particular concern was the self-paced nature of performance meaning that there was no guarantee that results would reflect the maximum performance capabilities of participants - they were simply coached and encouraged to choose a speed that was as fast as they could accurately perform. Clearly, participants may have differed in their interpretation of this, some perhaps performing more cautiously than others. Overall, this issue would not appear to have been a problem in terms of the grouped analyses, which involve a pooling of such variation. Indeed, the grouped analyses suggest that the methodology has generally been very successful, with results demonstrating meaningful and often subtle trends in response to the factors studied. However, the potential influence upon the individual participant analysis needs to be recognised. If future research is to focus more upon such analysis, as I have recommended, some refinements to the methodology are necessary. One particularly useful improvement would be to incorporate an element of pacing, something that would enable both a truer and more consistent measure of maximised performance to be achieved.

Concerning the general use of tempo as a measure of task skill, there has been nothing in this research that would suggest that it should not become one of the standard tools used by sight-reading researchers. Providing that suitable controls liked the ones just described are set in place, it would appear to offer a highly sensitive method of measuring performance ability, although there are obviously some methodological hurdles to be overcome before it can be used in the context of music with explicit timing variation.

The controlled preview methodology would have benefited from the recording of eye-movements to provide insight into the underlying pattern of perceptual activity. For example, it would have been useful to know whether the similar maximum effective preview spans of the two groups performing two-part material reflected the less skilled readers simply processing material more slowly, or whether different patterns of perceptual uptake were involved. The pattern-matching study would also have benefited from such a feature, and indeed one of the studies by Waters et *al.* (1997), upon which my own methodology was based, did include the facility. As well as again enabling possible differences in the underlying patterns of perception to be explored, this might have shed further light on the typical size of perceptual unit used. A further issue relating specifically to the pattern-matching study is that the process of remembering which was the correct key to press may have delayed the responses of some participants. Considering that the response times for non-matching pairs are not analysed, the issue could be resolved by employing a single key that participants press only when a match is perceived.

The motor sub-skill methodology was probably the least satisfactory within the thesis. As mentioned earlier, although it demonstrated an overall difference in motor skill between the two groups, the performances were too slow to provide any clear insight into normal sight-reading related motor activity. The evidence suggests that the slow tempi were in part the result of participants being overcautious due to the absence of any visual feedback, and so future implementations of this approach should incorporate a pacing mechanism to ensure that tempo is maximised.
The other main concern with the motor sub-skill methodology was the manner in which the experimental materials were displayed. It would seem that the use of finger numbers and key information made the task of memorisation more difficult for subjects than it might have been, and may have encouraged the use of representations not normally associated with sight-reading. Using an animation or video of the material to be performed had always been my preferred approach (see Chapter 11) but was not practicable for the current research. It is clear that such an improvement to the methodology is essential for future work of this nature.

13.4 Sight-reading pedagogy

There would seem to be a number of points worth mentioning in this area that have been raised by the research. With the evidence indicating that a variety of factors may have been responsible for hindering the sight-reading ability of my less skilled participants, the main teaching-related insight that can be drawn from the research in this thesis is that no single pedagogical regime is going to be appropriate to everyone's needs. To aid skill development, it would seem necessary to discover the specific limiting factor or factors in each individual case, and then to enlist appropriate practice strategies. Clearly this requires a considerable amount of further research into the 'anatomy' and development of skill, together with an investigation of techniques that might prove helpful in overcoming particular problems. Such work would necessarily involve longitudinal studies of many developing and experienced musicians carried out over an extended period of time.

Although the research findings cannot be used to support any specific pedagogical regimen for skill development, they do perhaps enable a more enlightened assessment of the potential usefulness of some methods. One particular case in point is the seemingly widely held view amongst instrumental teachers that less able sight-readers should be encouraged to look further ahead in the music. One traditional means of enforcing this stretching of preview has been for the teacher to direct perception to the desired area ahead of the notes currently being performed by manually covering over prior notation with a piece of card. Software implementations of this approach now exist (Souter, 2001). With my research indicating that skill difference is not primarily defined by variation in preview size but rather by efficiency of preview use, general use of this strategy would seem questionable. My findings suggest that a more appropriate path to skill development would be to use techniques encouraging the more effective processing of an individual's current preview range. It is possible, however, that the former approach might be useful if full preview capabilities are not being achieved due to an element of laziness in perception - or perhaps to further develop the abilities of already skilled readers.

A possibly useful generic teaching strategy is suggested by this thesis' novel use of tempo as a skill measure, an approach that enables the distinguishing of subtle differences in sight-reading ability. Sight-reading practice typically would seem to involve individuals performing music that is well beyond their zone of comfort. In preparation for graded examinations, for example, many students attempt to develop their abilities through practice tests appropriate to the grade level being taken, but are scarcely fluent with material from several grades below. It is possible, therefore, that many less skilled readers owe their lack of ability to having never built secure foundations and become fluent with easier music. Learning to read simpler music confidently beginning with a comfortable tempo, and then slowly increasing this over time using pacing technology, could provide an effective, evolutionary approach to skill development. With significant performance differences between the skill groups apparent even in the reading of monophonic sequences, it would seem wise to include even this most basic type of material within any regime of practice.

There has been a recent tendency amongst authors to be very upbeat about the universal trainability of sight-reading skill. For example, Thompson and Lehmann write that there is 'little evidence to suggest that "talent" has anything to do proficiency in sight-reading and improvising. Rather it is a case of diligent and inventive practicing (Thompson & Lehmann, 2004, p.154). Such a statement is not a little disingenuous, though - the extent of our knowledge on these matters is in reality too impoverished to be able to draw any reasoned conclusion. My research findings are not able to help in this regard; they have indicated a range of factors that might be responsible for hindering skill development, but can give no indication of how these factors come to be limiting to skill and the extent to which problems can be addressed through training. However, informal discussion with my less skilled reader participants revealed that a few had worked very hard at trying to improve their sight-reading ability, but clearly to little avail, at least in terms of achieving a skilled level of performance. Perhaps they had not yet found the appropriate key to unlock their potential, but the possibility must also be considered that their problems lay in factors that were not particularly conducive to major change. This may be because they were too old. For example, research by Lee (2004) suggests that sufficient practice prior to the age of 15 may be necessary for the development of skill. But it could also be that the typically dual-stave nature of piano sight-reading makes it a particularly demanding activity, one for which not all pianists have the individual make-up fundamental to developing proficiency. Proficient sight-reading on single melody line instruments would seem to be considerably less demanding and therefore perhaps more susceptible to training. Indeed, my results indicate that most of my less skilled readers would have been considered quite able readers in the context of performing normal monophonic music. They may have typically performed their experimental sequences more slowly than the skilled readers, but it must be borne in mind that the skilled group's performance of this material was very fast. Overall, until more exhaustive research has been undertaken, it would seem wise for researchers to keep an open mind on these matters.

13.5 Final comments

Given the dearth of research within the domain, the development of theory to explain why musicians differ in sight-reading ability necessarily involves a large element of uncertainty. A substantial empirical base is not just valuable as a stimulant to theoretical understanding; it also acts as a constraint upon it, making it less likely that there will be gaps in knowledge that researchers may be tempted to fill in with their own assumptions. As we have seen, the patterning account rests upon a number of such assumption-filled gaps, which my research has tested empirically, and found to be lacking in validity. The evidence indicates that the role of structure needs to be downplayed and that of other factors promoted. Although the multi-dimensional perspective suggested by my work seems to be considerably more evidence-based than the patterning account, caution is nonetheless called for. It appears unlikely that the patterning account can be resurrected, especially considering the findings of other research that is similarly questioning of it. But it would be wise for more study to be undertaken so that final conclusions can be more confidently drawn within a broader experimental context. Particularly, the studies in this thesis should be repeated with the proposed technological and methodological refinements, using different participants (perhaps younger, with less 'exposure' to musical structure) and a wider range and complexity of materials, especially involving more explicit stylistic elements. If the current findings are confirmed, research into sight-reading can then start to think more confidently in multi-dimensional terms, still sensitive to the influence of musical structure upon skill, but no longer constrained by the assumption that it is necessarily the primary factor. There is an enormous amount of further research that needs to be carried out, many of the required directions having been considered in this chapter and earlier in the thesis. The quantity of new information that my work would appear to have generated from just a small number of basic studies provides some indication of the rewards that a larger and more concerted research effort might achieve.

References

Abelson, R. P. (1979). Imagining the purpose of imagery. *Behavioral and Brain Sciences,* 2, 548-549.

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music, 23*(1), 3-16.

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50* (6), 1-80.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61,* 394-403.

Clarke, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209-236). London: Academic Press.

Clarke, E. F. (1995). Expression in performance: generativity, perception and semiosis. In J. Rink (Ed.), *The practice of performance: Studies of musical interpretation* (pp. 21-54). Cambridge: Cambridge University Press.

Connolly, C., & Williamon, A. (2004). Mental skills training. In A. Williamon (Ed.), *Musical excellence* (pp. 221-245). New York: Oxford University Press.

Elliott, C. A. (1982). The relationship among instrumental sight-reading ability and seven selected variables. *Journal of Research In Music Education, 30,* 5-14.

Ericsson, K. A., and Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245.

Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society of London, Series B, 266,* 2435-2440.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*(3), 219-272.

Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. *Visual Cognition, 10*(2), 201-232.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences 20*(1), 1-55.

Goolsby, T. W. (1994a). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception, 12*(1), 77-96.

Goolsby, T. W. (1994b). Profiles of processing: Eye movements during Sightreading. *Music Perception, 12*(1), 97-123.

Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31-50.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye-movement control. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 417-429.

Howell, D. C. (2002). *Statistical methods for psychology*. Pacific Grove: Duxbury.

Jacobsen, O. I. (1941). An analytical study of eye-movements in reading vocal and instrumental music. *Journal of Musicology, 3*, 3-32, 69-100, 133-164.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behavioral and Brain Sciences 17*(2), 187-245.

Kinsler, V., & Carpenter R. H. S. (1995). Saccadic eye movements while reading music. *Vision Research, 35*, 1447–58

Klemmer, E. T., & Frick, F. C. (1953). Assimilation of information from dot and matrix patterns. *Journal of Experimental Psychology, 45*, 15-19.

Kornicke, L. E. (1995). An exploratory study of individual difference variables in piano sight-reading achievement. *Quarterly Journal of Music Teaching and Learning, 6*(1), 56-79.

Kostka, M. J. (2000). The effects of error-detection practice on keyboard sight-reading achievement of undergraduate music majors. *Journal of Research in Music Education, 48*(2), 114-122.

Lee, J. I. (2004). *Component skills involved in sight reading music.* Peter Lang.

Legrand-Lestremau, S., Postal, V., & Charles, A. (2006). La vitesse de frappe est-elle liée au processus d'anticipation? *Le Travail Humain, 69*(1), 67-92.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12,* 182-195.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology, 15*, 1-29.

Lehmann, A. C., & McArthur V. (2002). Sight-Reading. In R. Parncutt & G. E. Macpherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 135-150). Oxford: Oxford University Press.

Lehmann, A. C., Sloboda, J. A., & Woody, R. H. (2007) *Psychology for musicians: Understanding and acquiring the skills.* Oxford: Oxford University Press.

Levin, H., & Kaplan, E. L. (1970). Grammatical structure and Reading. In H. Levin & J. P. Williams (Eds.): *Basic studies on reading.* New York: Basic Books,

Logan, G. D. (1982). On the ability to inhibit complex movement: A stop signal study of typewriting. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 778-792.

McConkie, G. W,. & Rayner, K. (1975). The span of effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578-586.

Mcpherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142-161.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner. K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during the sight-reading of music. *Current Directions in Psychological Science, 6*(2), 49-53.

Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin, 99*(3), 303-319.

Shaffer, L. H. (1976). Intention and performance. *Psychological Review, 83,* 373-393.

Shaffer, L. H. (1984). Timing in solo and duet performances. *Quarterly Journal of Experimental Psychology, 36*(A), 577-595.

Simon, H.A., & Chase, W.G. (1973). Skill in chess. *American Scientist, 61,* 394-403.

Sloboda, J. A. (1974). The eye-hand span: an approach to the study of sight reading. *Psychology of Music, 2* (2), 4-10.

Sloboda, J.A. (1976a). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28,* 1-16.

Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference of prose reading and music reading. *Canadian Journal of Psychology, 30,* 228-236.

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68*, 117-124.

Sloboda, J. A. (1978a). Perception of contour in music reading. *Perception, 6*, 323-331.

Sloboda, J. A. (1978b). The psychology of music reading. *Psychology of Music, 6*(2), 3-20.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35*, 377 – 396.

Sloboda, J. A. (1984). Experimental studies in music reading: A review. *Music Perception. 2*(2), 222-236.

Sloboda, J. A. (1985). *The musical mind.* New York: Oxford University Press.

Sloboda, J. A. (1991). Musical Expertise. In K. Anders Ericsson & Jacqui Smith (Eds.), *Towards a General Theory of Expertise* (pp.153-171), Cambridge: Cambridge University Press.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of fingering choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 185-203.

Souter, T. (2001). *Eye movement, memory and tempo in the sight-reading of keyboard music* (Doctoral Dissertation, Sydney Conservatory, 2001).

Strekfuss, R. J. (1984). The effect of a sight reading pacer machine upon the sight-reading ability of college wind instrumentalists (Doctoral Dissertation, The Catholic University of America, 1984).

Taylor, K. (1981). *Principles of piano technique and interpretation.* Novello.

Thompson, S., & Lehmann, A. C. (2004). Strategies for sight-reading and improvising music. In A. Williamson (Ed.), *Musical excellence* (pp.143–159). New York: Oxford University Press.

Truitt, F. E., Clifton, C. J., Pollatsek, A., & Rayner, K. (1997). The perceptual span and the eye-hand span in sight reading music. *Visual Cognition, 4*(2), 143-161.

Waters, A. J., Townsend., & Underwood, G. (1998a). Expertise in musical sight-reading: A study of pianists. *British Journal of Psychology, 89*, 123-149.

Waters, A. J., Townsend., & Underwood G. (1998b). Processing pitch and temporal structures in music reading: independent or interactive processing mechanisms? *European Journal of Cognitive Psychology, 11* (4), 531-553.

Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern matching paradigm. *Perception & Psychophysics, 59*(4), 477-488.

Weaver, H. E. (1943). A survey of visual processes in reading differently constructed musical selections. *Psychological Monographs, 55*(1), 1-30.

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research, 5*, 143-171.

Appendix 1

Controlled preview study sequences

C monophonic structured

F monophonic structured

G monophonic structured

C monophonic unstructured

F monophonic unstructured

G monophonic unstructured

C two-part structured

F two-part structured

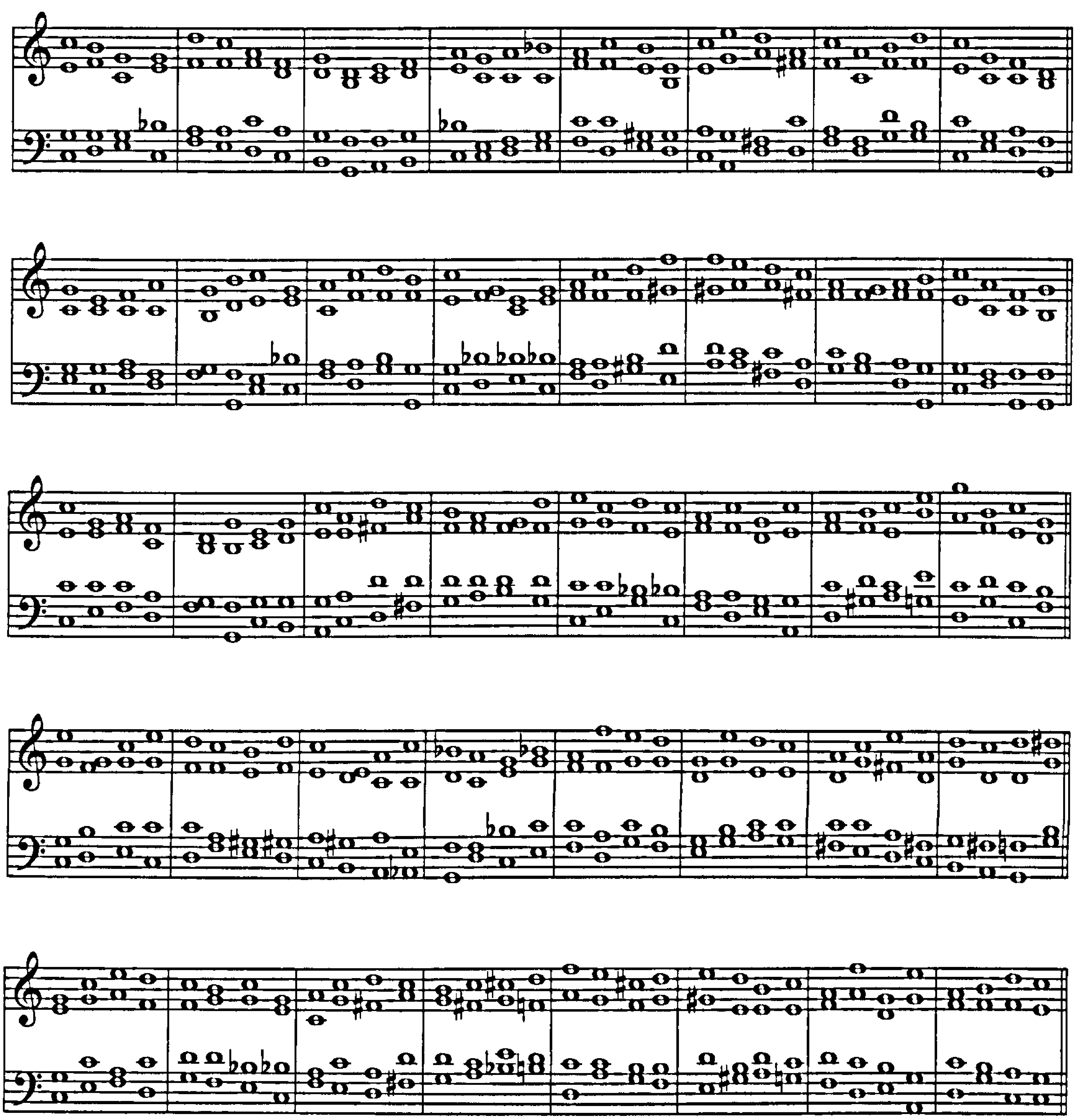
G two-part structured

C two-part unstructured

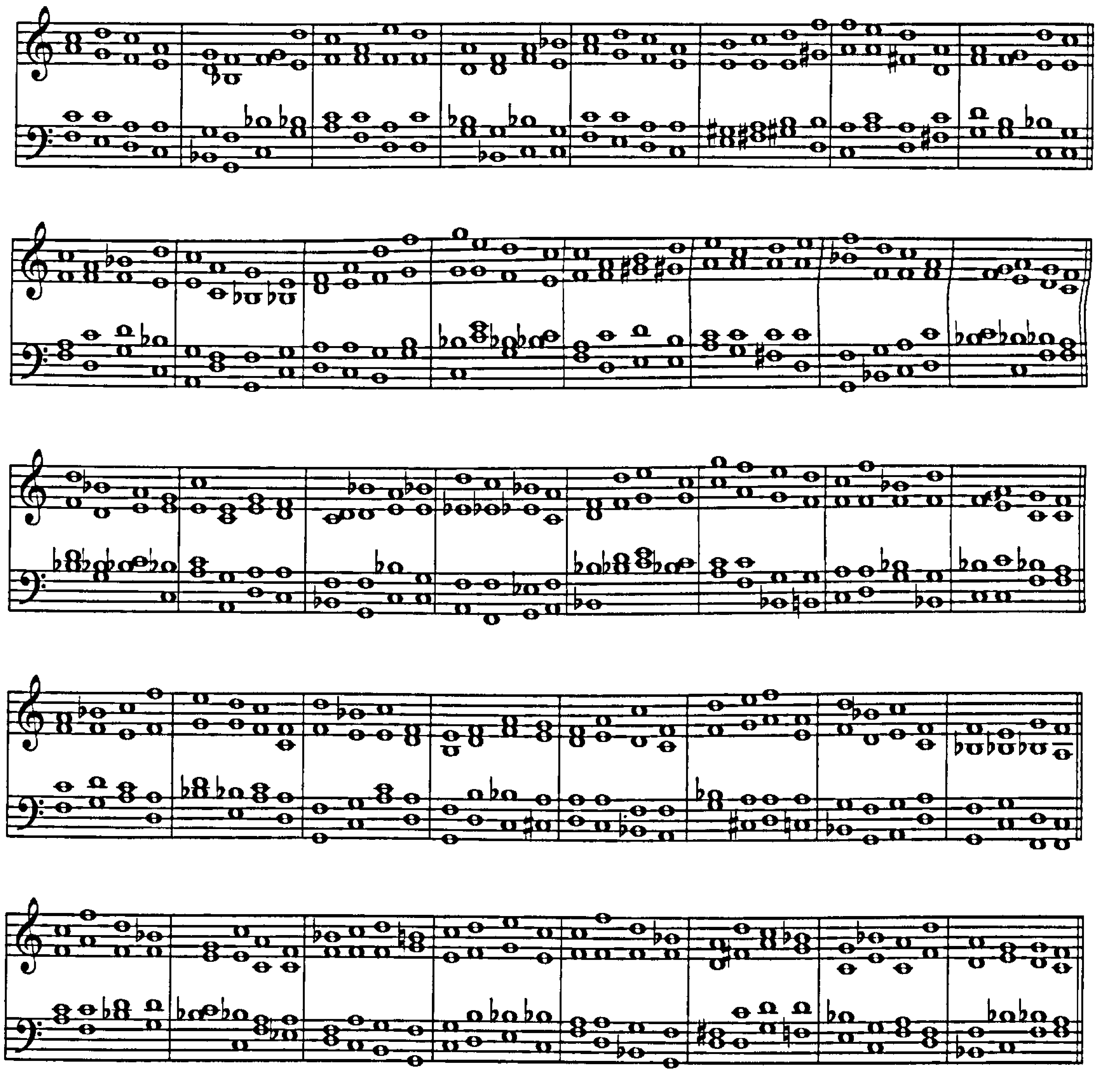
F two-part unstructured

G two-part unstructured

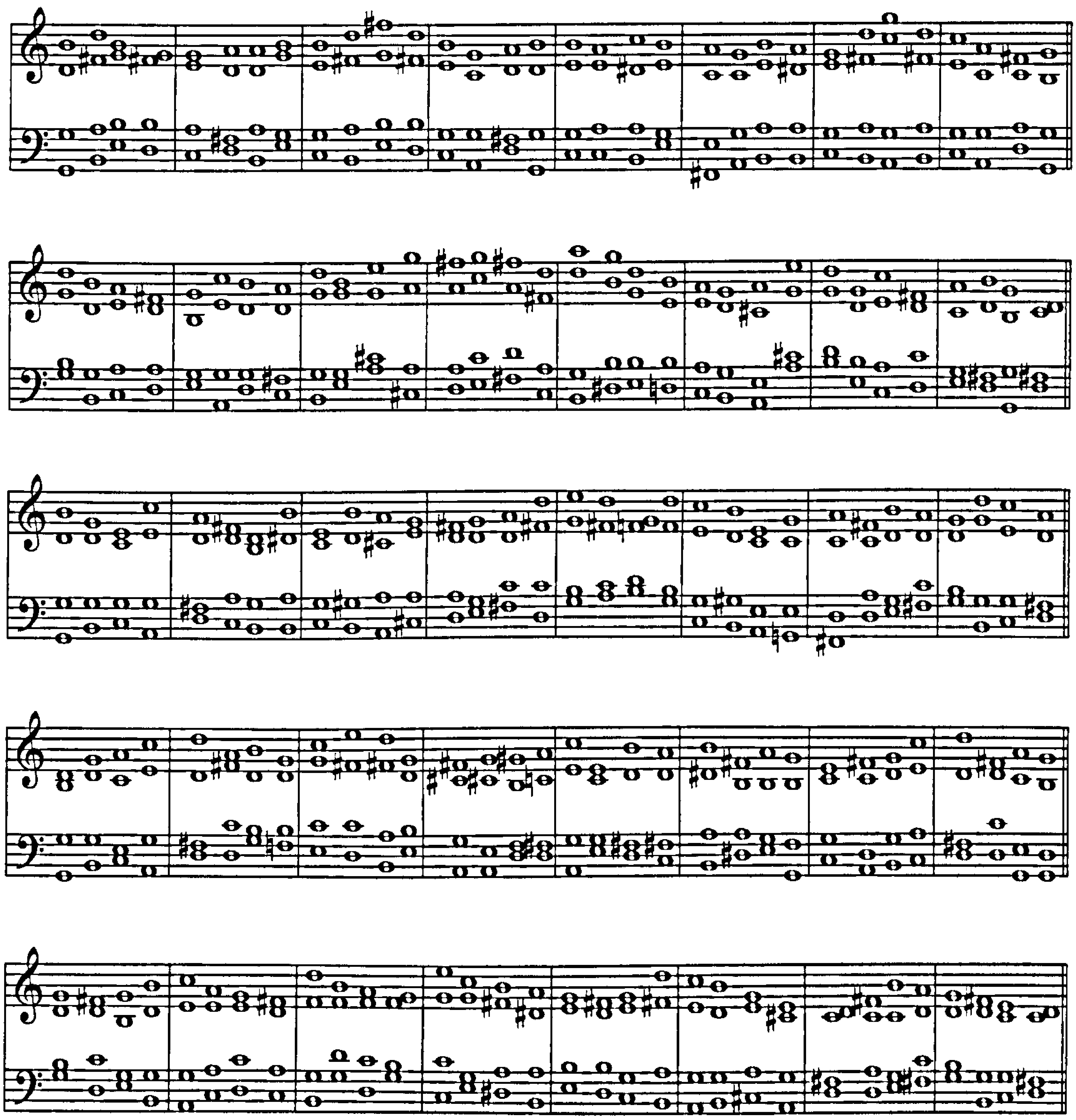
C four-part structured



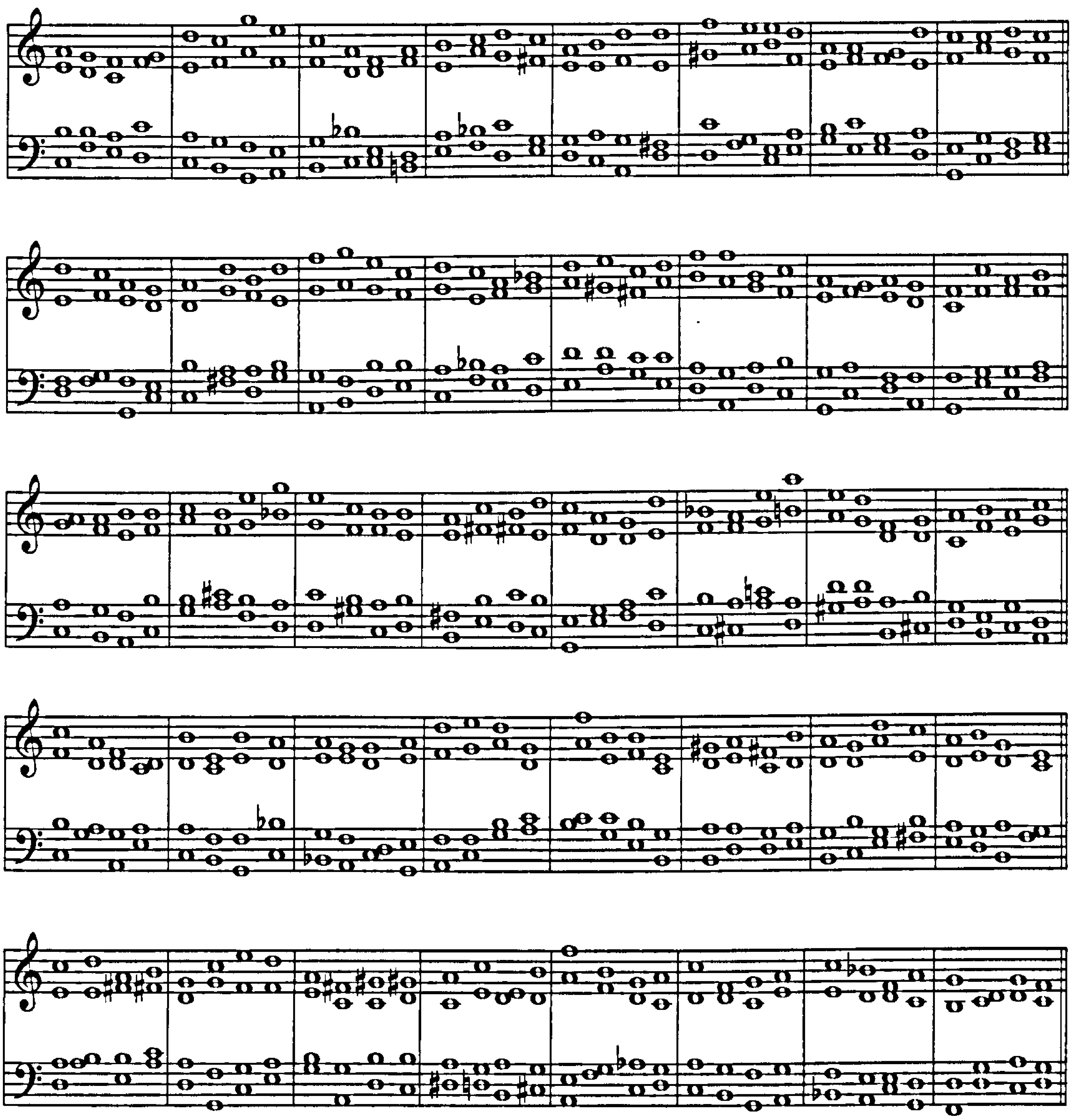
F four-part structured



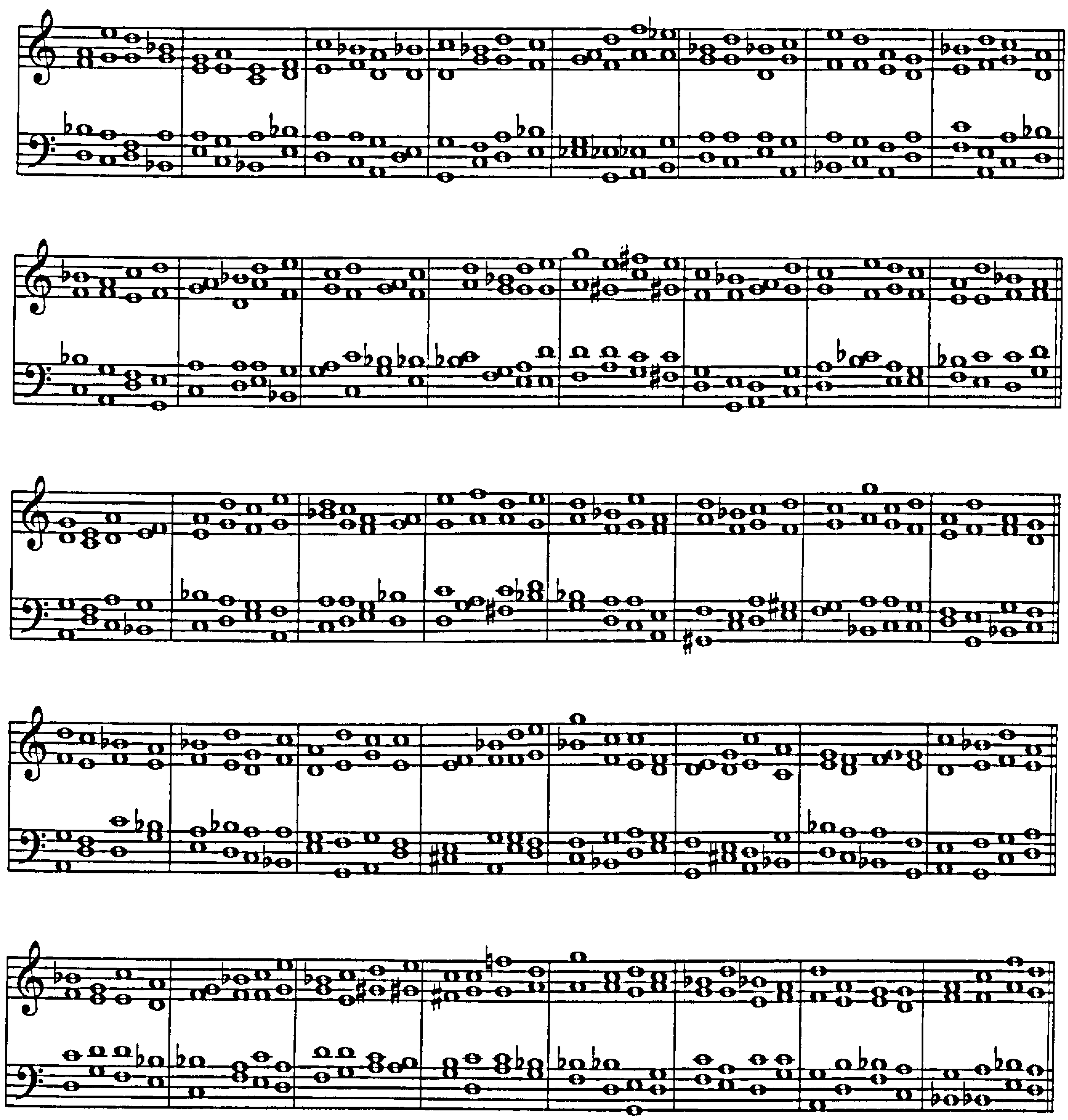
G four-part structured



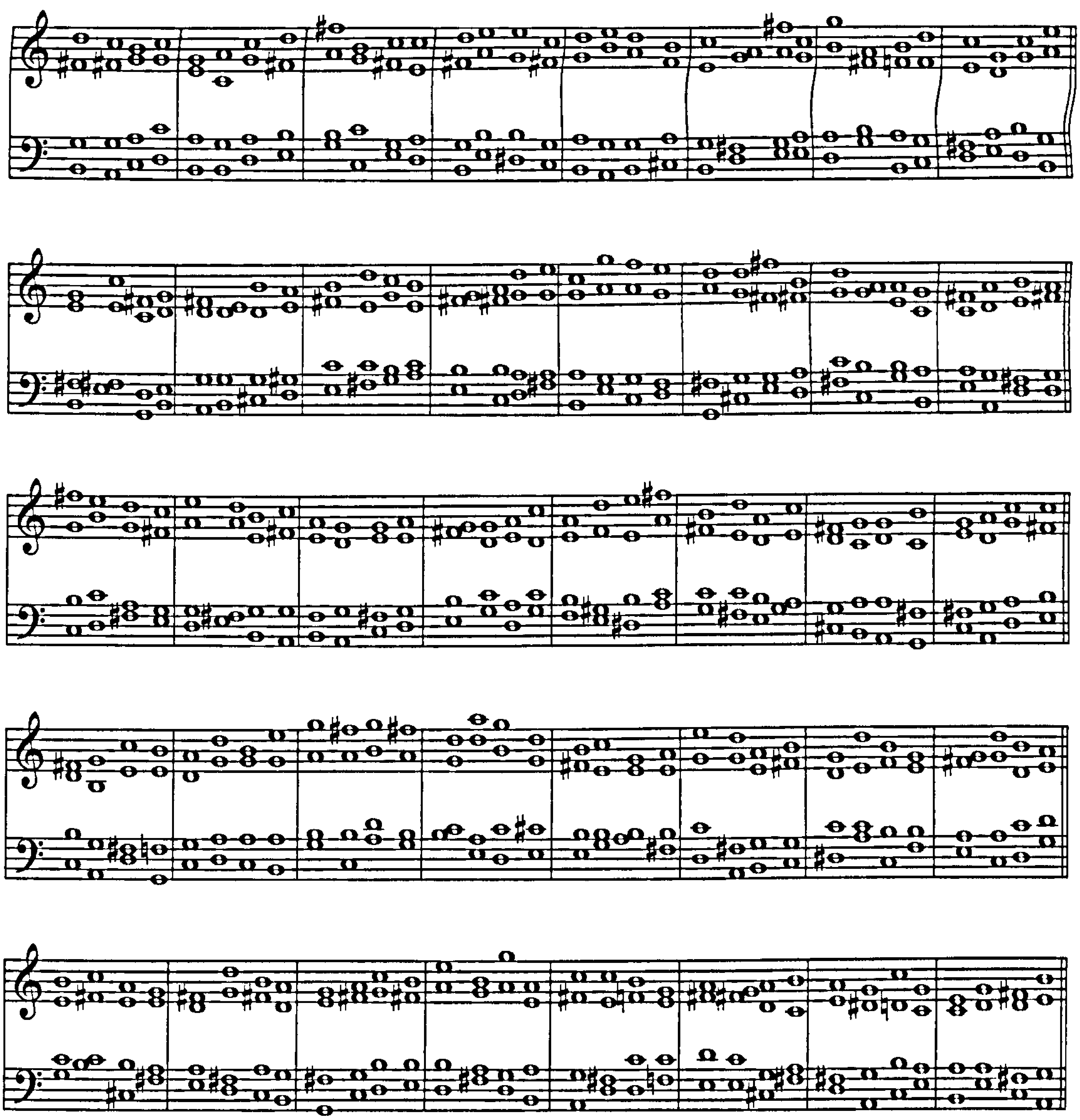
C four-part unstructured



F four-part unstructured



G four-part unstructured



Appendix 2

Perceptual sub-skill study sequences

1 note

2 notes

Structured Sequences

3 notes

4 notes

5 notes

Unstructured Sequences

3 notes

4 notes

5 notes

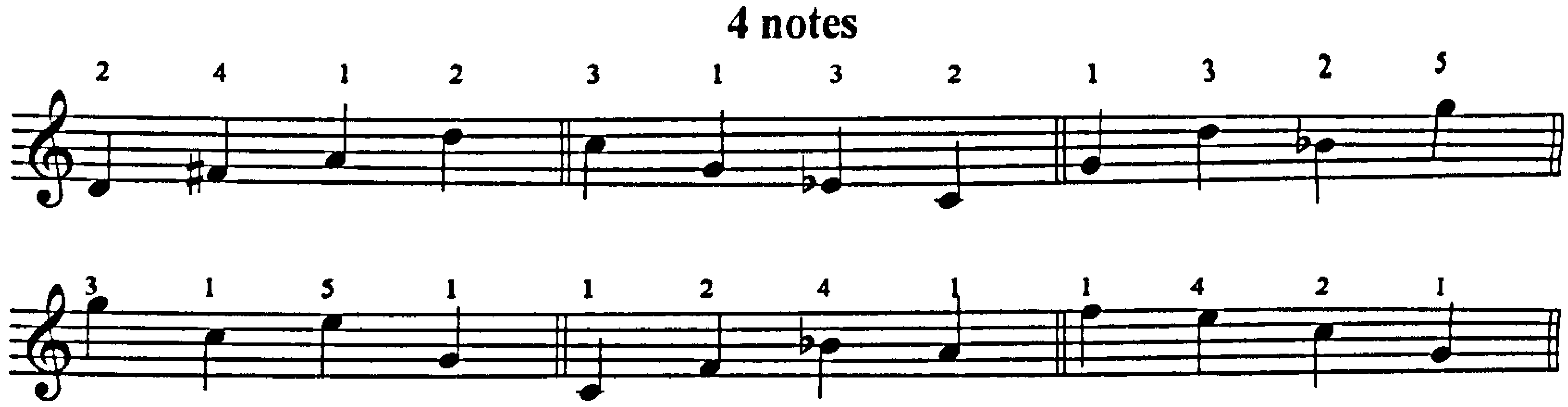
Appendix 3

Motor sub-skill study sequences

Structured Sequences

3 notes

4 notes

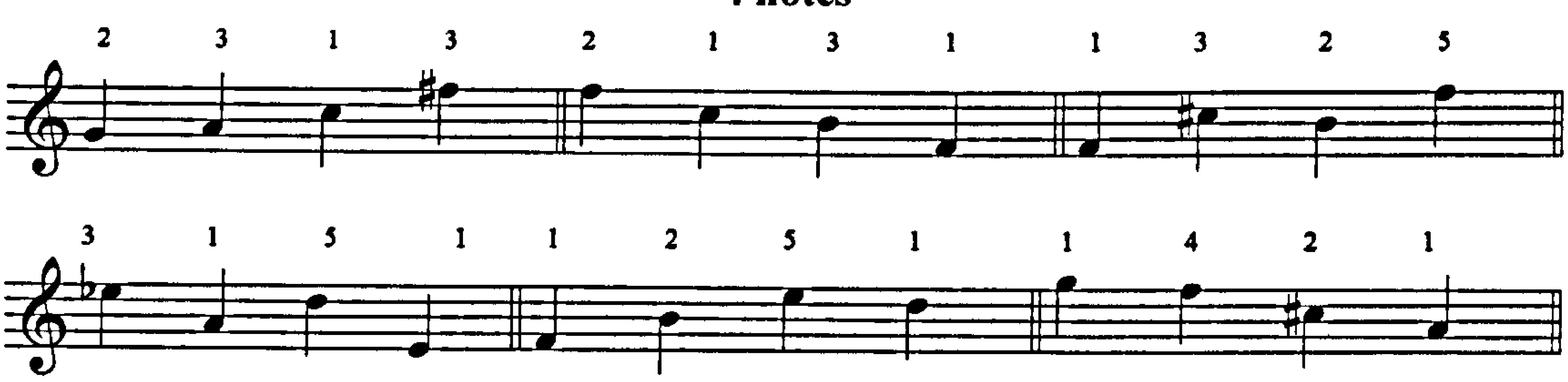


5 notes

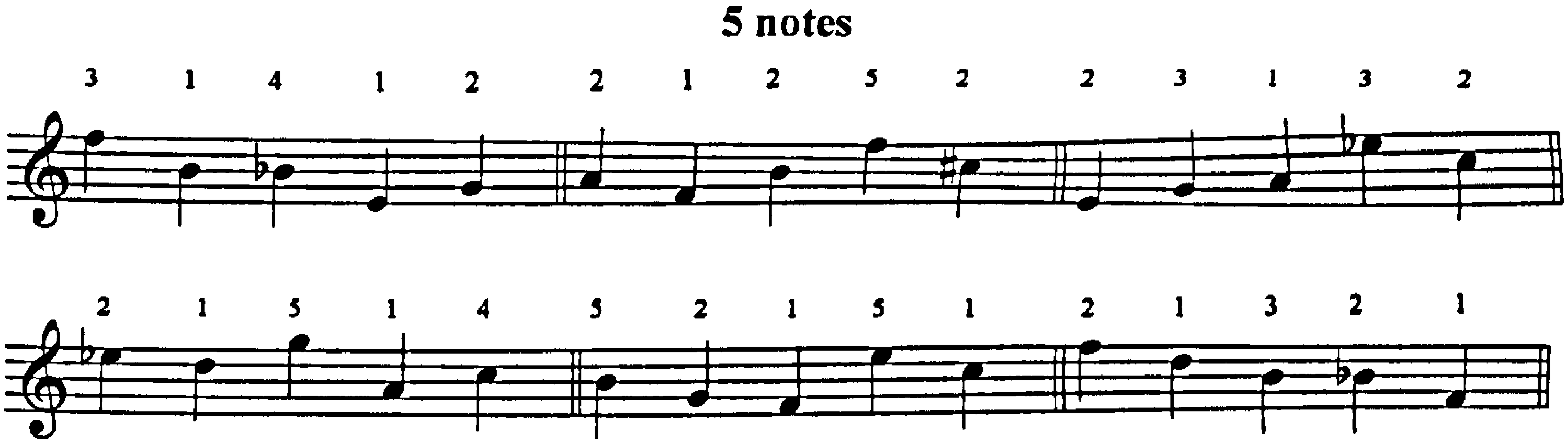
Normal unstructured sequences

3 notes

4 notes

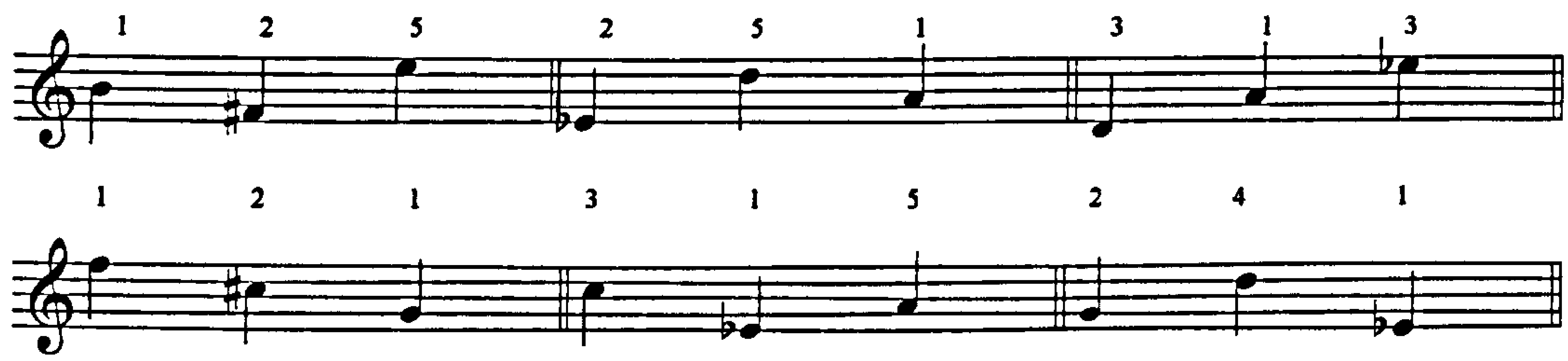


5 notes



Difficult unstructured sequences

3 notes



4 notes

5 notes