Ph.D in MUSIC EDUCATION

CHILDREN'S COGNITION OF TONAL ORGANISATION AS MEASURED BY REACTION TIME

RICHARD EDGAR HODGES

University of London Institute of Education

1992

Dissertation submitted in fulfilment for the requirements of the degree of Ph.D (Music Education)

ABSTRACT

This study examined perceptual and cognitive structures that children employ when listening to musical pitches. A number of experiments utilised reaction time as the dependent variable to identify perceptually salient factors in musical pitch perception, particularly the cognitive organisation of musical pitch in a tonal context. A chronometrically measured forced-choice paired–comparisons experimental paradigm was used with children between the ages of six and eleven, with the discrimination of *same* and *different* notes in context-free and various contextual presentations tested by a computer–driven environment.

Significant correlations suggest that the recognition of *same* and *different* notes in both context-free and contextual presentations was progressively facilitated, with responses exhibiting fewer errors and decreasing reaction times with increasing age.

Although no significant difference was observed in mean correct reaction times between uncontextualised *same* and *different* conditions, significant differences in reaction times were observed within each condition when suffix notes were each contextualised by a major triad prefix. Furthermore, while no significant correlation was observed between *same* and *different* notes in context-free presentation, the subsequent contextualisation by a major triad prefix to each comparison suffix note produced a significant positive correlation suggesting that the contextualisation effects were systematic.

A further experiment using a diminished triad prefix confirmed that the tonal specificity of the stimuli was related to the observed reaction times, with significant differences in correct reaction times for those stimuli which differed in the tonal range of their constituent pitches in relation to the circle of fifths.

The observed differences in the reaction time of responses were interpreted as differential measures of the internalisation of musical pitches to a cognitive structure such as a tonal schema. The hypothesis that perceptual facilitation of the coding of redundancy within such a recognised and practised cognitive structure such as tonality was supported for children of this age.

CONTENTS

3. METHODOLOGICAL CONSIDERATIONS AND THE DEVELOPMENT OF EXPERIMENTAL MATERIALS

4. PILOT STUDIES: *SAME* OR *DIFFERENT* RESPONSES TO TRICHORDAL NOTE-STRINGS IN DIATONIC CONTEXT

5. EXPERIMENT ONE: RESPONSES TO *SAME* OR *DIFFERENT* PAIRED-COMPARISON STIMULI

6. EXPERIMENT TWO: RESPONSES TO *SAME* OR *DIFFERENT* PAIRED-COMPARISON NOTES IN TRIADIC CONTEXT

7. EXPERIMENT THREE: TONAL SPECIFICITY AND REACTION TIME: *SAME* OR *DIFFERENT* RESPONSES TO PAIRED-COMPARISON NOTES IN DIMINISHED TRIAD CONTEXT

8. CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. STRUCTURAL RELATIONSHIPS IN MUSICAL PITCH PERCEPTION

1.1. Psychoacoustical and music-theoretic approaches

The perception of musical pitch has interested researchers for many years. Early studies of pitch perception were founded on the reductionist theory of Helmholtz (1863/1954), in which the perceived pitch of a complex note was attributed to the relative strength of the fundamental component. Under this theory, the higher harmonics of the harmonic series were considered to influence the timbre of a note but not the pitch. However, Schouten (1938) demonstrated that fundamental pitch may be perceived even when the fundamental is missing, confirming previous observations by Seebeck in the nineteenth century (Seebeck, 1841).

The purely psychophysical approach to perception of sound stimuli as exemplified by Helmholtz was subsequently rejected by the Gestalt psychologists, who proposed a number of laws of perceptual organisation which have had an important influence on present-day music psychology. Wertheimer (1923/1955) proposed that stimuli are grouped into configurations by various simple principles. For instance, the principle of *proximity* proposes that elements which are closer together are grouped more readily than those which are further apart. The principle of *similarity* states that like elements are perceived as similar. Another important principle is that of *good continuation,* which proposes that elements which follow each other can form perceptual groupings under certain conditions. The principle of *common fate* states that elements moving in the same direction are perceived together. The relevance of these grouping principles to music cognition has been examined by Deutsch (1982).

A number of musicologists (*e.g.* Berry, 1976; Meyer, 1956; Schenker, 1906/1954, 1935/1979; and Schoenberg, 1969, 1911/1978) have also proposed musical theories which have implications for the psychological processes involved in music perception. For example, the description of tonality by Berry (1976) implies an hierarchic structure of pitches that are perceptually non-equivalent in that one particular pitch serves as a cognitive reference point.

> *Tonality may be broadly conceived as a formal system in which pitch class content is perceived as functionally related to a specific pitch-class or pitch–class complex of resolution, often preestablished and conditioned, as a basis for structure at some understood level of perception.*
>
> *(Berry, 1976, p. 27)*

Empirical investigations by music psychologists of cognitive models based on music-theoretic principles have been largely concerned with Western tonal music, although some recent experiments with non-tonal and atonal materials have extended psychological theories of abstract internal representation.

Music-theoretic approaches have started analyses with acceptance of the basic principles of music theory. The sub-structures on which music is based, such as scales and triads, have been explored as representations of tonality. The use of scales and triads in experiments has given rise to much concern from certain investigators. For instance, Brown and Butler (1981) amusingly present the methodological dichotomy between psychoacousticians and musicians thus:

> *A psychophysicist might say:* *"If you allow too much music in the stimulus, or too much musical behaviour in the reporting task, you have too many variables. Control is lost; I can't tell which factors are and aren't operating."*
>
> *A musician might respond:* *"If you don't put enough music in the stimulus, and if you don't allow your subjects to act like musicians, then you aren't really telling me anything about music, and I'm just not interested in what you have to say."*
>
> *(Brown and Butler, 1981, p. 40)*

This recurrent problem of the constitution of both musical context and musical behaviour has affected the experimental work of music psychologists. Unfortunately, even the more recent cognitive-structural approaches to pitch perception that attempt to explain cognitive functioning are beset with similar problems.

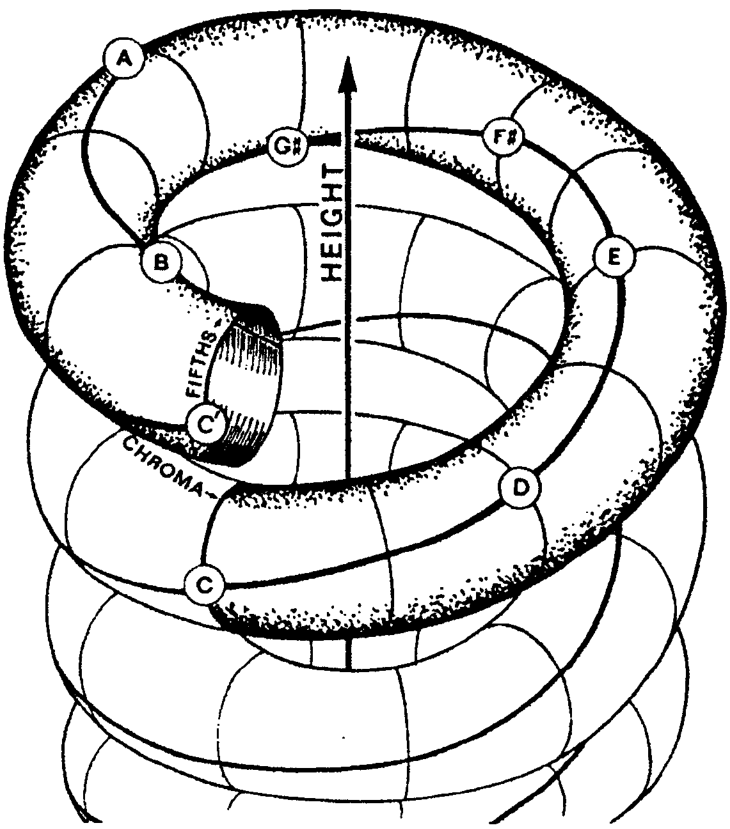
1.2. Cognitive-structural approaches

Examination by researchers of the relation of psychological theories of perception and cognition to musicological descriptions of musical structures is a relatively recent development in psychological research. Psychophysical investigations have developed unidimensional psychological scales of pitch. For instance, the *mel* scale interprets perceived pitch as a monotonically increasing function of the unidimensional scale of frequency (Stevens, Volkman, & Newman, 1937; Stevens & Volkman, 1940). However, a number of investigations have indicated that monotonic scales may be too simple to explain pitch perception. *Octave circularity* of relative pitch conceptualises pitch as bidimensional: *pitch height* is correlated with absolute frequency and *chroma* is correlated with relative position within an octave. Most bidimensional interpretations of pitch perception share the assumptions embodied in the model proposed by Drobisch (1846, cited in Ruckmick, 1929): pitch is conceptualised in a graphical representation as a helix, with pitch height represented by the vertical axis of the helix and chroma by the circular scale at its base, and with octaves located at corresponding points on successive turns of the helix. This bidimensional model was supported by Bachem (1950, 1954), who found that possessors of absolute pitch were consistently accurate in naming single notes but were often unable to classify their appropriate octave position.

Shepard's (1964) presentation of a cyclically repetitive sequence of complex notes composed of partials separated by octave intervals producing the illusion of an endlessly increasing sequence of pitch steps is the experiment most often cited as evidence for octave equivalence. Shepard (1964) provided a theoretical foundation for a structural approach to the perception of musical pitch. He recognised that the unitary conception of pitch as a direct analogue of frequency is inappropriate to explain the relationships perceived between pitches in music cognition. His interest in geometrical representations of cognition led him to formulate a complex multidimensional model, with musical pitches represented geometrically as a five-dimensional double-helix (Shepard, 1982a, 1982b). This geometrical representation (Figure 1.1) explores three components: a unidimensional projection of pitch height; a two-dimensional chroma-circle; and a two-dimensional representation of the circle of fifths. He argues that these relationships depict the perceived similarity of notes separated by small intervals such as the minor second, and the heightened similarity of notes separated by the octave, perfect fifth and perfect fourth.

Figure 1.1

Shepard's five-dimensional representation of musical pitch



(from Shepard, 1982b, p. 364)

This cognitive-structural model has been supported by a number of experiments. The most comprehensive empirical exploration of the cognitive representation of musical pitch has been undertaken by Krumhansl, whose studies have been most influential in determining the development of pitch perception research since her presentation of the influential tonal hierarchy model of perceived structure in 1979 (Krumhansl, 1979), which has been called the *tonal hierarchy theory* by commentators (*e.g.* Butler, 1989).

Krumhansl has criticised reductionist procedures as inappropriate (Krumhansl, 1983) since the perceptual or cognitive processes normally functioning during music listening may not be elicited or represented by the analysis of acoustical phenomena which are not embedded in musical context. Her studies purport to be concerned with the internal representation of musical stimuli and the processes involved in listening to music. Many of her studies use relatedness judgments in that subjects are required to judge how well one element follows another '*in a musical sense*' (Krumhansl, 1983, *p.* 35). The techniques of hierarchical clustering (Johnson, 1967) and multidimensional scaling (Kruskal, 1964a, 1964b; Shepard, 1962) have been used to graphically represent the psychological structure of such domains as tones in a tonal context (*e.g.* Krumhansl, 1979), chords in closely-related keys (*e.g.* Krumhansl, Bharucha and Kessler, 1982), chords in distantly-related keys (*e.g.* Bharucha and Krumhansl, 1983), and keys (*e.g.* Krumhansl and Kessler, 1982).

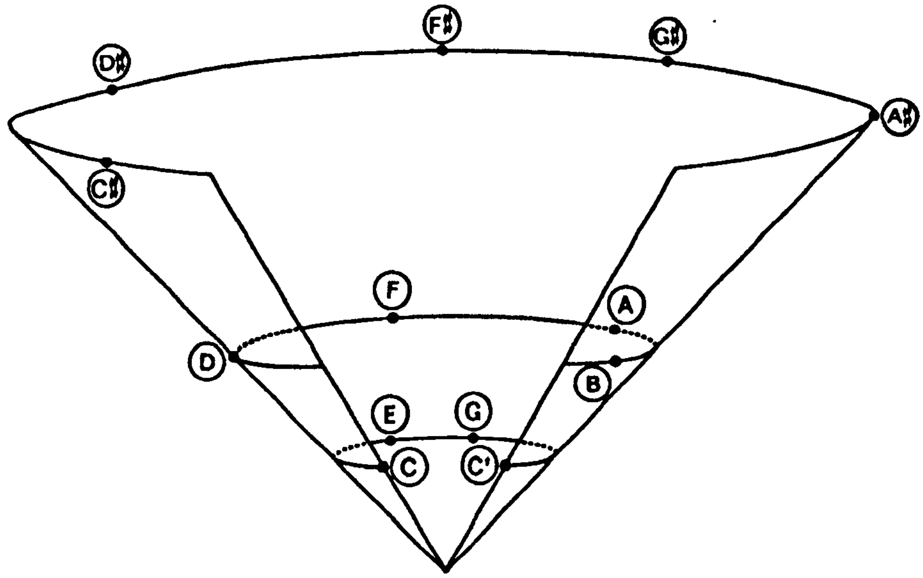
1.3. The tonal hierarchy theory

Krumhansl used similarity ratings to determine the cognitive representation of musical pitch in a tonal context (Krumhansl, 1979). She applied the notion of cognitive reference points (Rosch, 1975; Rosch & Mervis, 1975) to musical pitches, demonstrating that tonal organisation, the function of the relationship of the set of musical pitches to the tonic, is another important dimension in cognitive internal representation. Krumhansl presented pairs of notes in what she termed *'an explicitly tonal context'* (Krumhansl, 1979, *p.* 346) and asked listeners to judge how similar the first note is to the second note in relation to the tonal system suggested by the context. Judgments to three context types (*i.e.* a major triad chord, an ascending scale and a descending scale) were obtained using a seven–point response scale from very dissimilar to very similar. She found no significant differences between the context types in recovering similarity ratings.

Using the techniques of multidimensional scaling to represent the perceptual similarity of psychological relations, Krumhansl arrived experimentally at an idealised inverted conical representation of tonal relationships (Figure 1.2). The notes of the major triad are located on a plane near the vertex of the cone, the other diatonic notes somewhat farther from the vertex, and the non-diatonic notes furthest from the vertex.

Figure 1.2

Krumhansl's conical configuration of music pitch



(from Krumhansl, 1983, p. 40)

A number of problems are apparent in Krumhansl's otherwise elegant exposition of tonal relationships. The geometric representation presumes octave equivalence since the model exists within the tonal space of one octave. The *Shepard tones* used in her study attempted to negate effects of pitch height. These complex tones, whose partials consist of octaves of the fundamental, are passed through a bandpass filter that serves to keep average tone height constant regardless of fundamental frequency (Shepard, 1964). Krumhansl follows the earlier Shepard bidimensional model of musical pitch in adopting octave equivalence.

Krumhansl does not make explicit reference to the the experimental effect of *priming* caused by the continued use of the context defining stimuli in her experiments. It has been proposed (Dosher and Rosedale, 1989) that when a perceptual event is linked to a previous associative judgment (*i.e.* a *prime*) related concepts are residually activated which may help selection among close competitors when a stimulus is ambiguous.

Probably the the most problematic feature of Krumhansl's study, particularly for musicians, is the use of the term *atonal sequences*. Often the sequences used are quite tonal in implication, as one or two chromatic alterations outside a particular major scale suggest a different tonality rather than the complete absence of tonality, or confusion of tonal centres. For example the interpolated sequences from the second experiment of Krumhansl's study (Krumhansl, 1979) are constructed by raising the highest and lowest notes of the stimulus by a semitone:

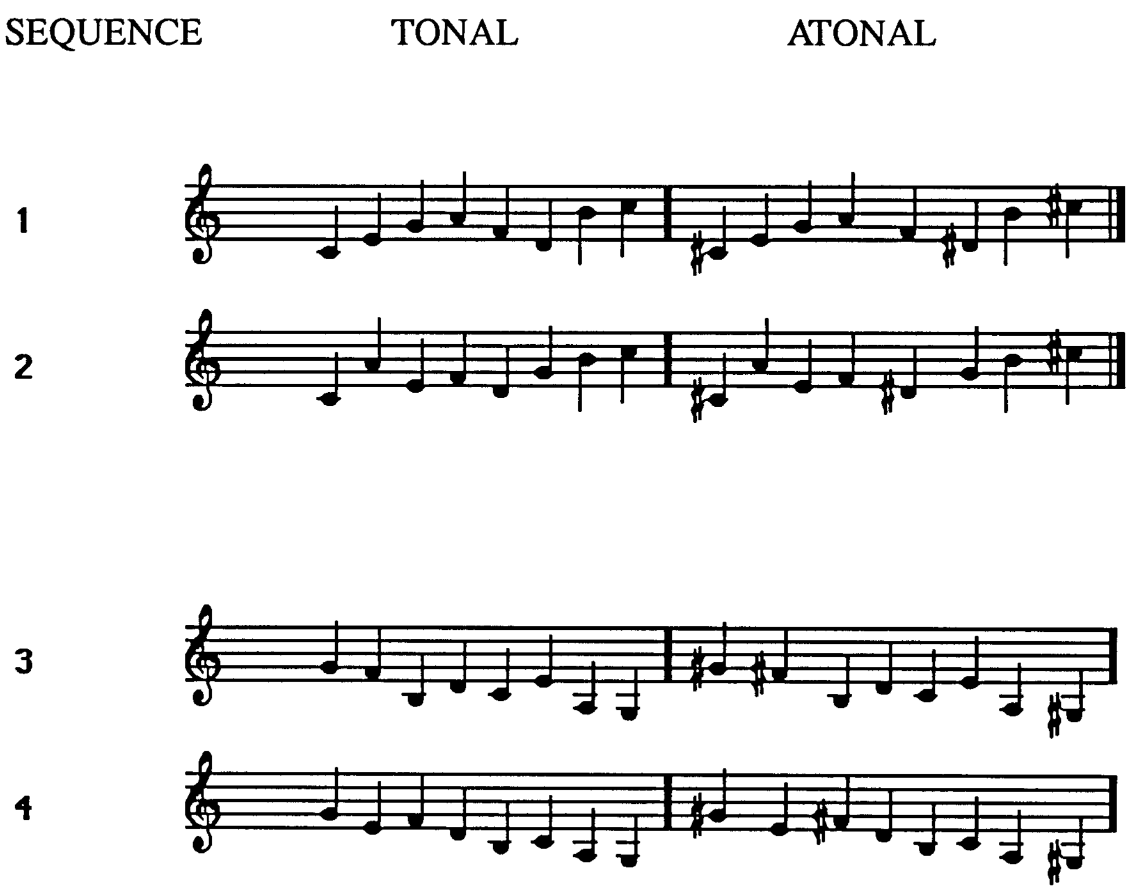
... thus destroying or weakening the tonality of the interpolated sequence.

(Krumhansl, 1979, p. 364)

An examination of some of Krumhansl's materials (Figure 1.3) shows clearly that numbers 3 and 4 of the so-called *atonal* sequences are not without tonal implication, as suggested by Krumhansl, as they may be interpreted in the tonality of *A minor*.

Figure 1.3

Test Materials



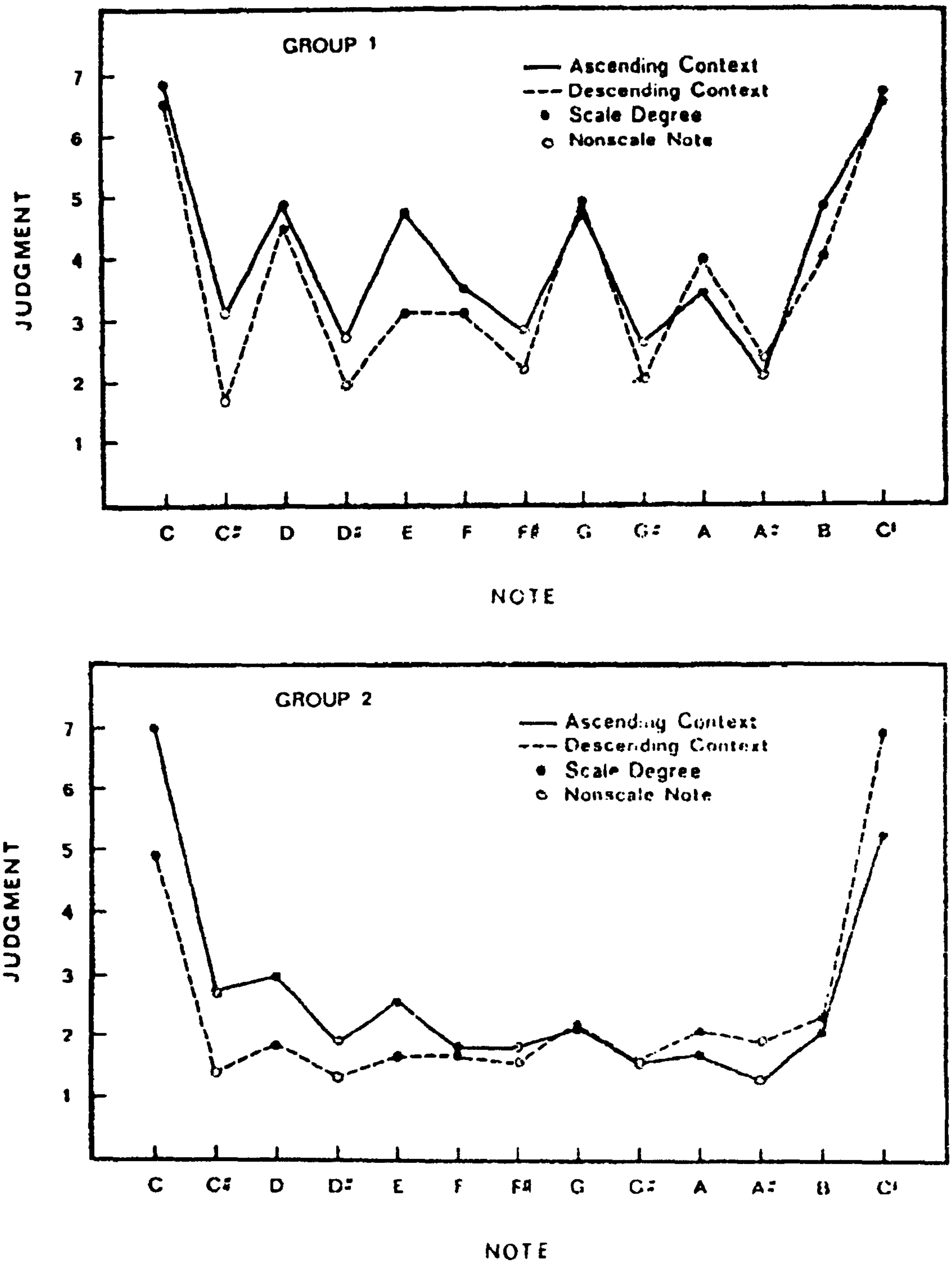
(from Krumhansl, 1979, p. 365)

Krumhansl found that diatonic notes were better remembered in tonal contexts and that nondiatonic notes were better remembered in her *atonal* contexts. It is surprising that Krumhansl did not check her assumptions concerning context by asking listeners to attribute a key–note to these stimuli. Krumhansl interprets her results as evidence for tonal contexts strengthening the representation in memory for diatonic notes and weakening the representation in memory for nondiatonic notes. An alternative explanation of the association between non–diatonic notes and atonal contexts might be related to the abstraction of the interpolated sequence to a tonal schema. Krumhansl's classification into tonal or atonal is simplistic at best and is not always consonant with the musical properties of the stimulus materials.

Krumhansl and Shepard (1979) attempted to quantify the hierarchy of tonal functions within a diatonic context. They developed the *probe–tone method*, in which a context–defining stimulus is followed by a single note, or probe–tone. Using a seven point rating scale, they asked listeners to rate how well each chromatic note within an octave range completed the context generated by a seven note ascending or descending major scale. They found that subjects with a moderate to high level of musical experience were more influenced by *octave equivalence* and the *tonal hierarchy* than less musical listeners. The hierarchy of tonal functions recovered by Krumhansl and Shepard from the eight listeners in the experienced musical group found the tonic note to be most preferred, followed by the other notes of the tonic triad, other notes from the diatonic scale, with the non-scale notes least preferred as completions to the context. This is represented graphically by the tonal hierarchy profile for Group 1 (Figure 1.4). The profile for the less musically experienced listeners shows that the average ratings for the octave are high, but that the other diatonic notes received lesser ratings of completeness (Group 2 of Figure 1.4). A third profile for the least musically experienced listeners of her sample recovered a curved distribution showing the effects of pitch height towards the octave. Krumhansl suggests that the differences in these three profiles show clearly the influence of previous musical experience.

Figure 1.4

Krumhansl Tonal Hierarchy Profile



(from Krumhansl and Shepard, 1979, p. 586)

Current research methodology offers many interesting and potentially useful research procedures. Hierarchical clustering and multidimensional scaling have been used frequently by Krumhansl and her associates. She has examined individual notes in tonal context as well as chords and keys in music-theoretic terms. Bharucha (1984) has examined anchoring effects in music, although his experimental materials could be considered to constitute a limited musical context. Certainly, the examination of pitch without reference to rhythmic factors has been questioned by Cross *et al.* (1991).

The establishment of the *tonal hierarchy theory* and the *probe–tone method* has resulted in a large number of extension studies investigating the psychological representation of chords and keys in a tonal context. Krumhansl & Schmuckler (1986) investigated the effect of the bitonal Petrushka chord. Bharucha & Stoeckig (1986, 1987) have looked at the priming of chords using reaction time measures. Krumhansl, Bharucha & Castellano (1982) have examined key distance effects on perceived harmonic structure. Krumhansl & Kessler (1982) have examined dynamic changes in tonal organisation in a spatial representation of musical keys. Krumhansl, Bharucha & Kessler (1982) have elaborated an interesting study of harmonic structure of chords in three related musical keys. Castellano *et al.* (1984) have investigated tonal hierarchies in the music of North India. All of these studies have found evidence for the perceptual reality of the tonal hierarchy theory.

Krumhansl and Shepard(1979) and more recently, Jordan (1987) and Jordan & Shepard (1987) have attempted examination of the tonal hierarchy by dividing the octave into more than twelve discrete steps and have concluded that the perception of microtones is strongly influenced by the tonal hierarchy. This, however, could be expected for musically experienced listeners. The recovery of the tonal hierarchy from a division of the octave into more than twelve discrete steps has not yet been undertaken with children. The division of the octave into twelve semitones has been arrived at by considering the interval of the fifth as the most important interval after the octave, but our harmonic system (at least those attributes of it adopted by music-theoretic psychologists) considers the third as the basis for triadic harmony (*cf.* Balzano, 1980).

Krumhansl has widened her terms of reference by considering the tonal hierarchy in relation to bitonal music (Krumhansl & Schmuckler, 1986) and twelve-tone serial music (Krumhansl, Sandell and Sergeant, 1987). However, the bitonal music study considers analysis in terms of simultaneous perception of more than one tonality. While there is no doubt that tonality as a structural principle in musical and perceptual organisation is of fundamental importance, it cannot be inferred that the tonal hierarchy as demonstrated by Krumhansl has general applicability to all musical styles and structures. What is clear is that any musical stimulus may be interpreted tonally, and attributed to a specific tonal centre. This happens in the perception of highly chromatic music, where harmonically ambiguous chords such as diminished sevenths may not be resolved and the implied tonality changes frequently. Such chords in music have tonal implications, although particular tonalities may be implied but not established.

1.4. Scalar conformance

Many of the experiments carried out by psychologists have investigated whether the structural characteristics of music are equivalent to cognitive processes which operate in music listening. The cognitive–structural models founded on music-theoretic principles have not escaped criticism from commentators. For instance, Shepard's formulation of a five-dimensional double helix (Shepard; 1964, 1982a, 1982b) has been criticised by both Cross *et. al.* (1991) and Hahn & Jones (1981).

One particular music–theoretic structure which has been investigated is the diatonic major scale. A number of studies have investigated scalar conformance. *Scale conformance* is determined by the extent to which certain configurations of notes occur within any one scale. A pattern of notes which does not occur in any one scale is considered as non–conformant.

Scale structure has been explored by Cross in a number of studies. Cross, Howell and West (1983) attempted to examine pitch–class sets and analyse scalar schema. They found that both musicians and non–musicians generally gave higher preference ratings of adjudged musicality to those melodies which were scale conformant. Howell, West & Cross (1984) experimented with the detection of notes incompatible with scalar structure. They found that the strength of scalar schema was determined by the relations between notes in the circle of fifths, with notes closer together (*e.g.* C,G,D,A) invoking a stronger schema than notes further apart (*e.g.* C,G,E,B). The investigation by Cross, West & Howell (1985) provides further evidence for the influence of the circle of fifths affecting judgments of the relation of suffix notes to various types of three–note prefixes. These studies of scalar relations extend the ideas of Brown and Butler (1981), discussed in the following section, concerning the unequivocal tonal implications of univalent trichordal strings of pitches.

Cross *et al.* (1983) have examined preferences for scale structure in melodic sequences. This study points out that Krumhansl's research methodology makes no distinction between tonal and modal presentations. Krumhansl (1979) plays her scale from C to C, for instance: what would the results be if the notes are re-ordered? For example, would listeners demonstrate the same set of tonal hierarchies if the same set of tonal relations in terms of intervals (C major diatonic scale) are presented as the Dorian mode (from D to D)? This distinction between scale structures and modal structures is something that has not been investigated. Furthermore, Cross *et al.* (1983) found that listeners imposed a rhythmic structure on the stimuli they presented and that this rhythmic structure determined the perceptual grouping of elements. Cross also investigated the importance of the circle of fifths in the determination of scalar structures.

Krumhansl's research has demonstrated the importance of training or developmental effects with a pronounced distinction between trained and untrained musicians (Krumhansl and Shepard, 1979). Cross's finding of no difference between musically trained and untrained listeners in his investigation of sense of *scalar conformance* contrasts profoundly with other research. Admittedly, Cross's samples were small and moreover consisted of homogeneous groups of adults. The absence of apparent training effects or differences between musically trained and untrained is puzzling: the results might indicate that sense of scalar conformance is something that is a fundamental aspect of music cognitive structures with minimal training influence, or that such sense of scalar conformance has very little at all to do with music cognition.

Whereas Krumhansl and her colleagues consider the tonal hierarchy as fixed and static (*i.e.* a set of relations from music theory applicable to certain music in particular), Cross *et al.* attempt to derive a relationship between the *circle of fifths* (with some ideas borrowed from Balzano, 1980) and sense of *scalar conformance.* For example, Howell *et al.* found that notes closer together in the circle of fifths (*e.g.* C, G, D, A) invoked a stronger schema than notes further apart (*e.g.* C, G, E, B).

Although the diatonic major scale has been proposed as a type of overlearned schema (Burns & Ward, 1982; Dowling, 1978) and attempts have been made by psychologists to investigate the extent to which prototypical scale structures might determine cognition (*e.g.* Krumhansl & Shepard, 1979; Krumhansl, 1979), this approach has elicited criticism from certain investigators, particularly Brown and Butler (1981).

1.5. Intervallic rivalry and the position finding theory

Brown and Butler (1981) find it difficult to agree with the demonstration of the tonal hierarchy as demonstrated by Krumhansl and Shepard (1979). They suggest that:

> *Using a scale to study tonality is analogous to using an alphabet to study grammar.*
> *(Brown and Butler, 1981, p. 44)*

More importantly, they contend that the study does not address higher–order musical structures in stimulus patterns, such as implied harmonic and tonal implications, or for that matter recognise that subjects bring experiences to these stimuli which enable them to be perceived in a higher-order musical context. However, such theoretical considerations have been discussed by certain investigators: Jones (1981, 1982), for example, has provided a framework for music perception in which *expectancy schemes* allow listeners to predict useful forthcoming notes that serve as cognitive attentional anchors termed *perceptual reference frames*.

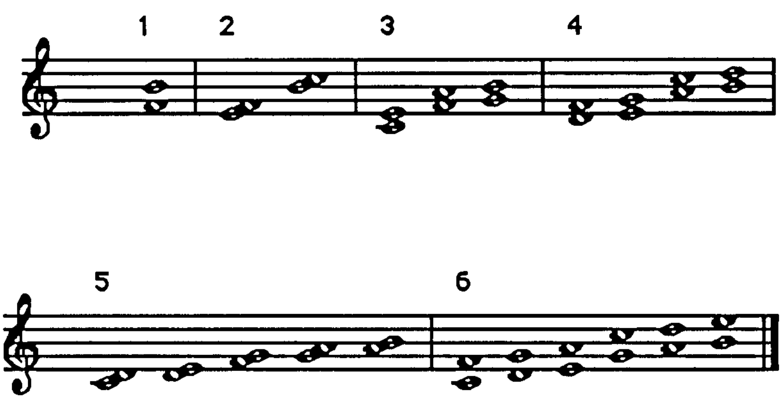
Brown (1985, 1987, 1988) found that a certain set of musical notes which suggest a particular key can be temporally re-ordered to suggest a different key. Her experiments used musical stimuli which were extracts from recognised musical masterpieces so that the pitches were embedded in a musical context. This approach contrasts with that of Krumhansl whose theory is without a temporal component. Krumhansl's model considers that notes are temporally invariant in their delimitation of a specified tonality, even though she accounts for modulation in some of her studies.

Butler (1989) proposed a theory of *intervallic rivalry* based on Brown's previous work. He suggests that it is the **rarest** occurring intervals in the diatonic set that give the clearest evidence of tonal implications. The rarest occurring intervals are the most important musical intervals from which inferences as to underlying tonality are formed. The tritone in particular, since it occurs in only one position in any given scale, is cited as the most tonally specific interval. The uniqueness of certain intervallic groupings as denoted by the number of occurrences of each configuration within a particular major diatonic framework is given in Figure 1.5.

Figure 1.5

The uniqueness of intervallic groupings within the major diatonic scale

number of occurrences:



(after Browne, 1981)

Each interval appears a different number of times in any given tonal set. This information can be used by the listener to determine the tonality of the music. The notion of intervallic rivalry dictates that listeners use the uniqueness of interval information to formulate the most likely candidate for the tonic. Browne (1981) proposes that the rarer intervals of the tritone and minor second, which occur once or twice in any particular tonality, provide enough information to categorise a particular tonal set.

The intervallic rivalry theory proposes that listeners use interval information to formulate the most likely tonic for a particular set of pitches. As certain intervals occur less frequently in any scale, these intervals are more important in formulating the tonal centre and are consequently more important in determining the tonality. Intervals therefore rival each other in terms of importance in establishing tonality.

Butler (1990) has extended his intervallic rivalry theory in examining post-tonal music. One of his most important contributions to the understanding of the perception of tonality is to have shown that the tonal perception of pitch intervals is time dependent. In other words, any perceived hierarchy is affected by note order relations and is not invariant. Other investigators (*e.g.* Palmer and Krumhansl; 1987a, 1987b) have also shown that temporal considerations affect perception and these commentators have acknowledged that the presentation order of pitches and rhythmic characteristics can modify the context-defining properties of note groups.

1.6. Recovery of the tonal hierarchy with children

Krumhansl & Keil (1982) examined children's acquisition of the hierarchy of tonal functions in music. They reported increased differentiation with age between tonic triad and other diatonic notes, and early internalisation of key structure. Their study of the hierarchy of tonal stabilities of children of elementary school age suggests a developmental sequence, showing increased differentiation of musical notes as established by a diatonic melodic context. The first developmental feature to emerge concerns the distinction between diatonic and non-diatonic notes (grades 1 and 2) followed some years later (grades 3 aged 4) by the tonic triad versus the other scale components. Older children (grades 5 and 6) in Krumhansl and Keil's study showed a preference for a note from the tonic triad in either of the final two positions of the musical stimulus. Only adults showed a strong sensitivity to octave equivalence, which would support the notion that pitch height is a more salient perceptual characteristic for children. If the hierarchy appears developmentally as suggested by Krumhansl and Keil (1982), it is reasonable to suppose that processing takes place along these dimensions. If this is the case, then the ability to classify scale and non–scale notes will appear at an earlier age than the ability to classify elements such as tonic triad, other scale components, octave equivalence or temporal asymmetries of note order.

However, this pattern contrasts with the finding of Krumhansl's earlier research (Krumhansl and Shepard, 1979) involving training effects in adults. That study suggested that musical experience produces a different sequence of development: pitch height is followed by octave equivalence and the diatonic–nondiatonic distinction is followed by a growing awareness of the tonic triad versus other scale notes. However, training effects with adults are not necessarily comparable to the developmental components of children's cognition. The importance of octave equivalence to adults demonstrated by Krumhansl and Keil (1982) may indicate that developmental and training acquisition sequences are distinct. Krumhansl (1983, *p*. 42) suggests that further work is needed to clarify the acquisition order for different aspects of tonal organisation.

However, Dowling (1982) argues that children and adults both show the same developmental sequence of melodic information processing, *viz.* pitch, contour, tonality and interval size. He maintains that training enhances tonal scale structure and proposes that a scalar schema is a fundamental component of the cognitive capability of all listeners:

> *... one effect of training is to enhance the importance of the tonal scale system in information processing of melodies. The intervals of the scale system are firmly embedded in the minds of even untrained listeners. Training facilitates the application of that system to new materials.*
>
> *(Dowling, 1982, p. 427)*

Krumhansl's methodology has been used subsequently by others (*e.g.* Speer & Adams, 1985; Speer & Meeks, 1985). Speer & Adams (1985) examined the tonal hierarchy theory and found very substantial differences in performance between musically trained and untrained children. However, they examined their data only in relation to the tonal hierarchy, and did not comment on some interesting features which are present in their data, particularly the development of octave equivalence suggested by their findings. Their results indicate that the development of pitch perception is highly trainable and not controlled by the operation of some domain-independent cognitive growth principle:

> *The existence of such a powerful training effect seems inconsistent with the existence of some intact perceptual structure for music. Our result is not, of course, inconsistent with an interactive model, in which experience instantiates open variables in some pre-existing structure.*
>
> *(Speer & Adams, 1985, p. 15)*

Speer & Meeks (1985) found that second grade children made no distinction between diatonic and non–diatonic notes when the context was an ascending scale, whereas Krumhansl and Keil's study (1982) observed this characteristic in their subjects. This inconsistency may be attributed to differences between the stimuli. Krumhansl and Keil used a two–note suffix to a four note triadic context (*i.e.* C, E, C, G). Speer and Meeks, on the other hand, replicated the single note suffix to a seven–note scalic context (both ascending and descending). Since the probe-tone methodology from which a rating profile is established has been adopted by other investigators with little modification, it seems important to determine to what extent test materials themselves may have contributed to obtained results.

1.7. Misunderstandings of musical structure in psychological studies

Music-theoretic psychological research is founded on ideas and assumptions from music theory. Certain writers (*e.g.* Cross *et al.*, 1983; Hahn & Jones, 1981; Kallman, 1982) have suggested that this is an inadequate way to proceed since musical stimuli and experimental design may inadvertently determine subject's responses.

Trehub *et al.* (1986), Badertscher (1985) and Cuddy and Badertscher (1987) have unreservedly accepted the premises inherent in Krumhansl's demonstration of the tonal hierarchy in their studies with children, sometimes betraying a limited understanding of musical structure. These misunderstandings found in psychological research based on music–theoretic principles are to be expected to a certain extent as much published work is undertaken by psychologists who have limited experience and training in music. Some of the important American psychomusicologists (*e.g.* Deutsch, Balzano, and Krumhansl) have influenced other psychologists who have not always fully understood music–theoretic principles.

Trehub's investigation of semitone discrimination with infants (Trehub *et al.*, 1986) is an example of misunderstanding of musical structure. She adopts stimuli based on Krumhansl and Keil's (1982) experiment but apparently fails to appreciate that the semitone distinction which she is examining falls outside diatonic structure as the stimulus is part of an augmented triad. The use of an augmented triad of C, E and G# to discuss diatonic structure in C major when this chord does not appear in any major key is puzzling. A further problem with many of the studies which follow Krumhansl's probe-tone methodology is that no distinction is made between major and minor tonalities and the chordal structures which characterise each specification of a tonality. According to the intervallic rivalry theory, an augmented triad of C, E, and G# is more tonally specific of the tonality of *A minor* than a simple C major triad is representative of the tonality of *C major.* The augmented triad may not represent such a stable structure as a major triad, since a discord, in musical terms, requires resolution. However, the augmented triad of C, E, and G# is perfectly acceptable in the tonality of *A minor*, and is still diatonic, although Trehub might think otherwise. More problematic is the use of five-note sequences from which a whole set of generalisations about children's cognition of music are proposed. Trehub herself (1987) has pointed to some of the deficiencies of her research.

More recent research seems to acknowledge that atonal melodies are difficult to generate for test purposes. Morrongiello and Roes (1990) examined the effects of musical training on developmental changes in children's perception of musical sequences by matching line drawings to 9-note melodies. They found that children aged 9 years performed better than children aged 5. However, they define an *atonal* melody as:

> *... one that is often said to be out of key, in other words, non–diatonic notes are included.*
>
> *(Morrongiello and Roes, 1990, p. 814)*

However, examination of some of the experimental stimuli reveals an atonal melody as C-D#-F#-A#-C-A#-F#-D#-C. While this is clearly outside the tonality of C major, the aural perception does not even fit the description non-diatonic. The four different pitches (as C,Eb,Gb,Bb) are unequivocally in Db major or Bb minor in that they constitute a seventh chord within either tonality.

Many of the melodies used by experimenters as examples of atonality are tonally specific in that they include a context–delimiting tritone. This misunderstanding of musical structure betrays a rudimentary understanding of music theory, and inappropriate classification into tonal and atonal (or nondiatonic) has marred certain experiments. If Butler's conception of intervallic rivalry underlies perceptual processing of the cognition of tonality, then the results of Morrongiello and Roes (1990) are puzzling.

1.8. Inconsistencies between the theories of tonal hierarchy and intervallic rivalry

Krumhansl's demonstration of the tonal hierarchy (Krumhansl, 1979) involves two assumptions which have questionable psychological validity. The tonal hierarchy theory assumes that the cognitive reference point (tonic) is determined unequivocally from the stimulus and that the tonal hierarchy evidenced is an abstraction to an already learned set and is not just a product of the context materials used. More importantly, the theory proposes that such a model of relations has psychological reality for listeners engaged in musical cognition. Krumhansl (1990a) has acknowledged that context-creating properties of stimuli are distinct in that cognitive structures established by scales may be slightly different in tonal strength from those invoked by triads and chords. This helps to explain observed differences between studies, but questions the psychological validity of the method.

The notion of cognitive reference points (Rosch, 1975) seems crucial to Krumhansl's explanation of tonal hierarchy observed in preferences demonstrated by subjects. However, it does not seem unreasonable to argue that stimuli such as the diatonic scale are not really bringing into play a cognitive structure used in music listening but that the context of the experiment instantiates a set of expectancies, or schema.

A fundamental difference between Krumhansl's *tonal hierarchy* and Butler's *intervallic rivalry* is the role of the tonic, or reference note, in relation to other pitches. For Krumhansl, all pitches heard are related to the tonic as a reference note. This means that subjects presented with a stimulus such as the three different pitches of the tonic triad (*cf.* Krumhansl and Keil, 1982) are abstracting a reference note from those pitches. While this may be the case, certain selections of notes would not unambiguously specify a particular scale. For example, the notes E, F and G belong to a number of different scales (C major, D minor, F major, F minor, *etc.*) and could therefore be compatible with a number of different tonics. This pattern of semitone followed by a tone occurs twice in the major scale (starting on the mediant or leading note) and twice in the minor scale (beginning on either the supertonic or the leading note).

This inconsistency between the two theories concerning the role of the tonic results from the assumptions underlying Krumhansl's experiment. The tonal hierarchy demands relation to a single note, the base of Krumhansl's conical representation (*cf.* Fig 1.2) and therefore it might be argued that such a demonstration of the tonal hierarchy is in part a result of the experimental methodology itself, even though the tonality is determined by scrutiny of the subject's probe tone responses. Any musical stimulus is conceptualised within a specific tonal centre, and the relationship between notes is fixed with a presumed tonic at the base of the cone. However, intervallic rivalry does not demand that the listener chooses the note to which other notes are to be related as the tonal hierarchy would suggest, but that certain reference notes are eliminated as a melody or tune progresses. This is quite different. Such a theory presumes the schema to operate on an exclusive basis whereby rejection of possible tonics establishes the best scalar representation which can accommodate the tonality of a stimulus. Krumhansl implies that notes are heard in relation to an internally held reference point, but Butler suggests that notes are heard in relation to each other, related to the notion of scalar conformance. In other words, internalised abstraction to a pitch set must be a fluid, constantly changing representation.

The tonal hierarchy theory has received much attention from investigators following Krumhansl's seminal work. The methodology of similarity rating scales linked to probe-tone techniques has been the usual demonstration of the psychological reality of the tonal hierarchy. Certain replicatory studies have acknowledged the dynamic nature of the psychological representation of chords and keys in a tonal context. For instance, Krumhansl, Bharucha & Kessler's (1982) study of harmonic structure of chords in three related musical keys has examined the dynamic fluctuation of musical cognition. Krumhansl and Castellano (1983) have also acknowledged the dynamic nature of music perception. However, the tonal hierarchy theory has elicited criticism from a number of investigators, notably Cook (1987a) and Butler (1989, 1990).

Cook (1987a) in a review of the journal Music Perception has criticised Krumhansl's methodology and other similar psychological research in that:

> *... the results obtained from such experiments are a function of the contextual properties of the particular stimulus selected, and cannot validly be generalised to a given diatonic set per se.*
>
> *(Cook, 1987a)*

This argument may be valid for the early demonstrations of the tonal hierarchy. However, it can reasonably be argued that the stimuli fairly represent the structures employed in music that do pertain to a diatonic set. Cook has problems with Krumhansl's concept of '*musical sense*' (Krumhansl, 1983, *p*. 35) as he considers that the linear and melodic structure of musical materials is as important as diatonicism implied by or abstracted from test materials.

Brown's research has come closer to an understanding of the relationship of specified note combinations to tonality perception by using musical stimuli which have a temporal component. However, Butler's theory of intervallic rivalry, derived from Brown's previous work, seems inconsistent with evidence from experiments with children which have explored the tonal hierarchy theory (e.g. Badertscher, 1985; Cuddy & Badertscher, 1987).

The findings of Cuddy and Badertscher (1987) do not seem consonant with the intervallic rivalry theory. Their data do not support the notion that a position-finding mechanism is used by children in their perception. They found that the diminished triad (which contains the tritone) was a poor context-defining stimulus whereas the tonic triad provided a much clearer context.

This study also contains a number of findings which are unexpected under the tonal hierarchy theory. In two related experiments, Cuddy and Badertscher asked children and adults to rate on a seven-point scale how well a probe–tone provided a musical completion to each of three contextual patterns (*i.e.* tonic triad, major scale and diminished triad in C major). They found that children asked to judge between E, F, G, or A as completions to the major triadic context of C major rated these pitches equally acceptable, whereas higher ratings for either E or G could have been expected under the tonal hierarchy theory.

Cuddy and Badertscher also found that adults rated F higher than E for the triadic context, and rated both A and B higher than E following the major scale context. This high rating for F can be explained as F is the best candidate for the tonic if C–E–G is interpreted as the the dominant chord of a perfect cadence. This is not suggested by Cuddy and Badertscher as an explanation, and is an example of the problem encountered by many psychomusicologists who fail to appreciate the dynamic nature of functional diatonic progressions.

Probably their most revealing observation was that the note C received high ratings when it followed the major scale and major triad contextual prefixes, but failed to induce a high rating following a diminished triad context. However, Cuddy and Badertscher do not point out that whereas the note C appears twice in both scale and major triad prefixes, it does not appear at all in the diminished triad context (which has two leading notes). This suggests short–term memory effects influenced by the particular properties of the stimulus characteristics themselves, rather than any higher order perceptual processes.

Although the intervallic rivalry theory may well explain the behaviour of experienced music listeners, it seems inconsistent with the evidence provided by studies with children using the probe-tone methodology, such as that from the study by Cuddy and Badertscher (1987). Clearly, what is needed to investigate the cognitive processes of children is a study using a different methodology.

1.9. Implications for experimental design

It is proposed here that tonality is not only a feature of music itself, but a cognitive system of abstract mental representation which is utilised to make sense of the complex phenomena of music. Young children's abilities to abstract musical pitches to a tonal schema may not be as well developed as those of adults. It has been suggested that children perceive music in a different way from adults (*e.g.* Speer and Meeks, 1985), although the processing differences have not been described in detail. Cognitive processing differences would certainly help to explain developmental differences noted by Speer and Adams (1985).

No researcher has yet tested the position finding theory with children. Such a test of this theory with children would lead to a greater understanding of the perceptual and cognitive processes involved in perception. However, the methodology previously used with adults locating a tonic from a given stimulus might not readily be understood by children. Young children would be unlikely to verbalise the tonal relations between pitches since the explicit description of tonal structure requires training.

A different methodology to clarify children's perceptual processes is required. A methodology which does not use similarity rating scales would be appropriate. The research problem lies in constructing a measurement of children's perceptual processes that can be elicited without recourse to some kind of classification or discriminatory response requiring terminology with which children are unfamiliar. In other words, the problem concerns the nature of the output that can be measured with children as young as five or six. A comparison choice between stimuli might be appropriate, but a rating scale would be better than a simplistic choice between, for instance, *same* or *different* musical stimuli. It is how different such stimuli are perceived to be which is important, but this information is difficult to elicit with young children.

The primary research question would concern how children mentally represent musical pitches in memory. This leads to a consideration of whether they assimilate pitches to a scalar schema as suggested by the tonal hierarchy theory or whether they use rarely occurring intervals as postulated in the position finding theory. It is possible that either or both strategies could be used in certain circumstances.

The mental abstraction of a musical stimulus and its assimilation to a scalar schema must take time, however brief, to process. If it is postulated that certain groups of notes will more easily be assimilated than others, then it is reasonable to propose that the mental processing required to respond would be quicker for some groups of notes than others. This would lead to differential reaction times being observed in some circumstances to a classification response.

In attempting to determine if processing of melodic materials is hierarchical, the importance of ascertaining exactly along which dimensions such hierarchical processing takes place is paramount.

If differences in processing time are observed for stimuli in a choice classification task, it seems reasonable to suppose that hierarchical processing of some kind may be operating and that abstraction to a schema must be instantiated. This assumes that greater processing time will be required for deeper nested levels within the hierarchy. An experimental procedure, therefore, which critically examines and quantifies the times taken for *same* or *different* responses with musical materials exhibiting different degrees of scalar conformance must be the obvious starting point for empirical investigation.

The most fruitful methodological solution for the present study would be to utilise the time taken to classify particular stimuli according to the criterion of *same* or *different* as a basis for the measurement of cognitive processing. This indication of processing as measured by response times might throw some light on developmental issues. The methodological approaches to reaction time responses are considered in the review of reaction time experiments in the next chapter.

2. REACTION TIME MEASURES AND COGNITION

2.1. Introduction

This chapter begins by outlining the models of information processing which have been suggested by data from reaction time measures. This is followed by a summary discussion of the experiments which have been concerned with perceptual matching, particularly choice reaction time experiments. The binary choice response between *same* or *different* is a discrimination task which children should be able to undertake with musical stimuli. Studies with children which use probe tone methodologies as developed by Krumhansl (*e.g.* Krumhansl and Keil, 1982) or pitch predominance (*e.g.* Temko, 1971) are problematic in that children are required to understand what constitutes what Krumhansl has termed *'musical sense'* (Krumhansl, 1983, p. 35) in relation to tonality. Discussion of reaction time experiments with musical stimuli follows, beginning with chronometric studies of the musical interval sense by Balzano (1977, 1982). The application of stage reduction theory to music cognition, notably by Fiske (1982a, 1982b, 1985, 1987, 1990) is followed by a consideration of other studies of music cognition which have used reaction time measures.

2.2. Information Processing Models

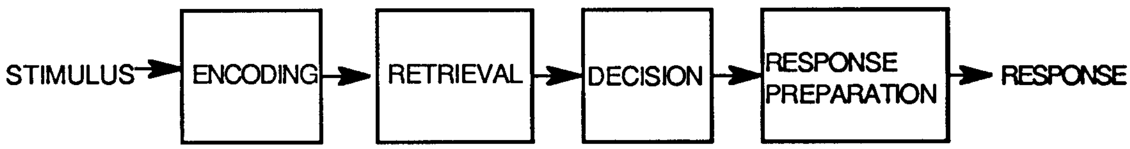
People do not respond instantaneously to incoming sensory information. Time is required for the mental and physical processes that precede a consequent overt response. The measurement of the time separating presentation of stimulus and observation of response, the reaction time (RT), has occupied psychologists for many years. Much research has concerned the nature of the processing mechanism, and whether serial or parallel processing appropriately describes observed behaviour.

2.2.1. Serial processing

The *discrete stage model* of Sternberg (1969) represents a serial model of cognitive processing. Each stage in the process is considered to be chronologically dependent on successful completion of the previous stage. At the end of each processing stage, the total output is passed to the next stage, and no stage has access to the partial products of the previous stage. An idealised diagrammatic representation of the discrete stage model (Figure 2.1) makes clear this division into distinct functional stages. Under this model, an RT may be interpreted as equating the summed durations of all the stages. Adoption of this model allows the absolute duration of mental processes to be estimated since they are considered to be separate. For instance, the insertion of an additional process to a task should allow the duration of the additional process to be measured by subtracting the mean RT of the shorter task from the RT of the longer task.

Figure 2.1

Diagrammatic representation of the *discrete stage model*



reaction time → → → → → → → →

(after Sternberg, 1969)

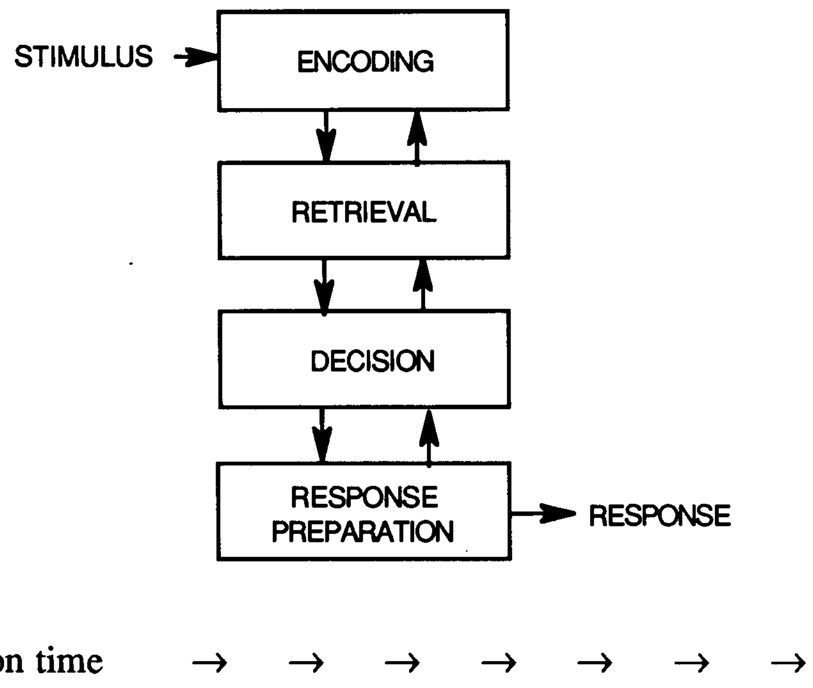
The arrowed lines in this representation indicate the direction of the products of each stage, and indicate clearly the unidirectional nature of the flow of information. Each stage is consequently dependent on the previous stage and the total processing time undertaken is the combination of each processing stage.

2.2.2. Parallel processing

An alternative explanation of cognitive processes is presented in parallel processing models. The *cascade model* (McClelland, 1979) postulates the division into functional stages, but the operations occur concurrently and information flows continuously in the form of an increasing spread of activation. An idealised diagrammatic representation of the cascade model is presented in Figure 2.2.

Figure 2.2

Diagrammatic representation of the *cascade model*



(after McClelland, 1979)

This parallel model shows clearly that the stages are concurrent. The freedom of information exchange is indicated by the bidirectional arrowed lines. This indicates that each subsequent process has access to the partial products of the previous stage, which is not the case in the serial model.

The interpretation of the cascade model is problematic as the absolute duration of component mental processes cannot be ascertained by simple subtraction. As the processes are not strictly successive, mean RT would not necessarily equal the summed durations of all the processes. However, although parallel processing may be more difficult to interpret than serial processing since the stages are not distinct, additive effects on mean RT may still be meaningfully interpreted if significant differences between RTs are observed.

2.2.3. Non-redundant hierarchical processing

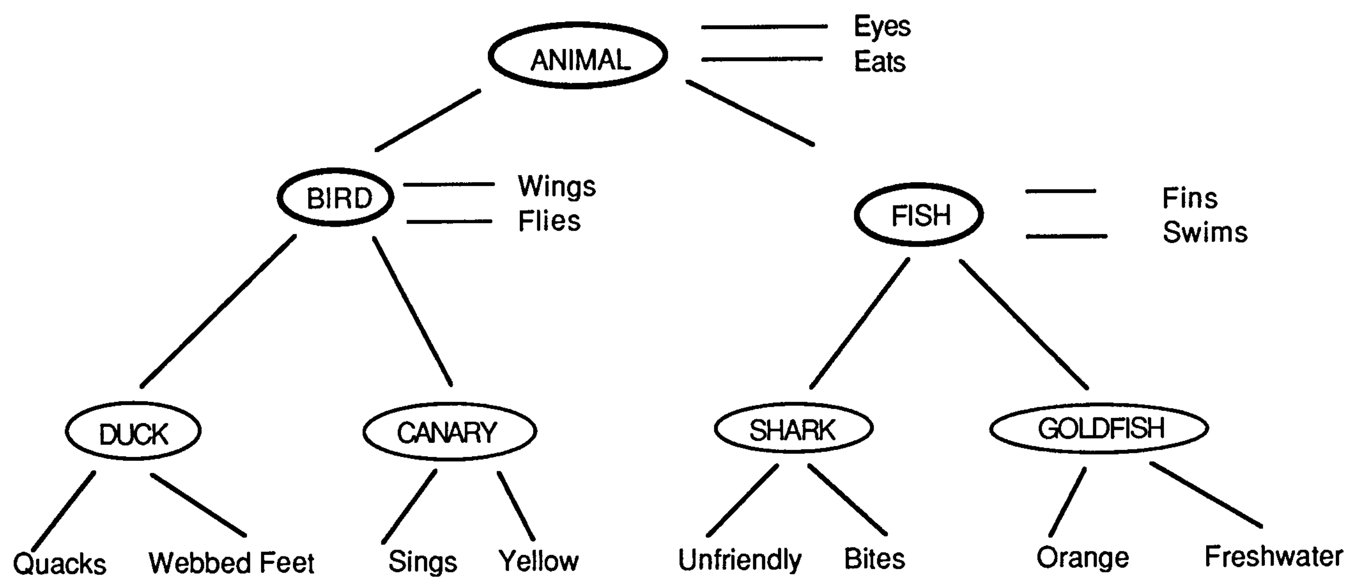
Empirical studies concerning the priming of conceptual frameworks, or perceptual reference frames, have frequently been undertaken for lexical decision judgments, *i.e.* the recognition of words.

A classic RT experiment by Collins and Quillian (1969) measured subject's responses to such questions as "Does a duck quack?" and "Does a duck have eyes?". They found that questions concerning commonalities between animals (*e.g.* eyes) had longer RTs than specifics related to a particular animal (*e.g.* duck and quack, canary and sings). They interpreted this to suggest that memory organisation is best modelled as an efficient and non–redundant hierarchical structure. The model is termed non–redundant as each aspect of categorical information is stored only once at the highest appropriate level within the hierarchy. Those comparisons which are closer together consequently take less time to process. Each cognitive unit or concept within this representation is linked to only one other concept and retrieval involves a spread of activation from each unit to adjacent units or nodes. This example is represented as in Figure 2.3.

This model of Collins and Quillian (1969) represents the *spreading activation* theory which is currently considered to be the best description of priming of associative memory judgments. The model is hypothesised as nodes (*i.e.* concepts) which are linked (*i.e.* associated) to each other. Priming is seen as the activation of particular nodes with connected concepts spreading throughout the nodal network.

Figure 2.3

Semantic memory non–redundant hierarchical model



(after Collins and Quillian, 1969)

2.2.4. Age differences

An important factor in RT experiments is the effect of age differences on the speed of cognitive responses. Much research has established that children process more slowly than adults (Sternberg and Rifkin; 1979). Bisanz *et al.* (1979), for example, asked subjects of various ages to determine whether pairs of pictures were identical visually or in name. Subjects judged the name of the pictures more slowly than the visual similarity, and this time difference was used to estimate the processing time required to retrieve the name from memory. RTs decreased with increasing age (*i.e.* 8 year olds responded to the names of common objects in 282 milliseconds; whereas 10, 12 and 19 year olds response times were 210, 142 and 115 milliseconds respectively). This effect has been found in other studies. For example, Kail (1988) tested subjects between the ages of 8 to 22 on a visual search task and a mental rotation task. He found evidence for increasing speed of response with increasing age in the sample of 8 to 22 year olds, described well by an exponential function with a common rate of change. He proposed that a general mechanism affects processing speed which changes with age.

A complication with RT experiments concerns the relationship of processing time and psychomotor response mechanisms. The changes in processing time related to age differences which have been noted by some investigators may not be merely concerned with dealing with stimuli but may also reflect the time taken by subjects to arrive at an operational output. The time taken to respond to stimuli may not be attributed wholly to cognitive processes, since some kind of operational delay is inevitable.

Some studies have examined auditory perception across various ages. For example, Galton (1899), investigated the time required by subjects of various ages to press a key immediately on the presentation of a sound stimulus. However, there appear to be no studies examining the developmental changes of cognition by using RT responses to musical stimuli with children or with adults.

A further problem of interpretation is the distinction between cognitive speed and cognitive development suggested by decreasing RT measures. The practice effect of repeatedly undertaking a task improves efficiency and there is some evidence to suggest that repetition of item familiarity results in decreased RTs. For example, Kristofferson (1972a) found that RTs with a fixed set of alphabetic character classification items decreased as a result of practice. However, such practice effects have been largely negated when items have been changed between trials (*e.g.* Kristofferson; 1972b).

2.3. Perceptual Matching and Reaction Time Measures

A perceptual matching task requires matching one stimulus with another. This process is a fundamental component of information processing. A number of experimenters have utilised the forced–choice *same* or *different* paradigm with RT measures.

2.3.1. Bamber's two-process model

Bamber (1969) investigated RTs required to recognise whether two successively presented rows of letters were the *same* or *different*. He was concerned to test the four possible types of model which had been suggested by previous research (cf. Egeth, 1966). Processing is undertaken along one dimension at a time in *serial* models, and processing along a number of dimensions occurs simultaneously in *parallel* models. For example, the comparison of two stimuli which could vary along the dimensions of colour, shape, and tilt of an exterior line were used by Egeth (1966). Such stimuli are known as *multidimensional*. Within each of the serial and parallel models, processing can also be either *self–terminating* or *exhaustive*, depending on whether a *different* response is generated immediately a difference is recognised in one dimension or whether all dimensions have to be processed.

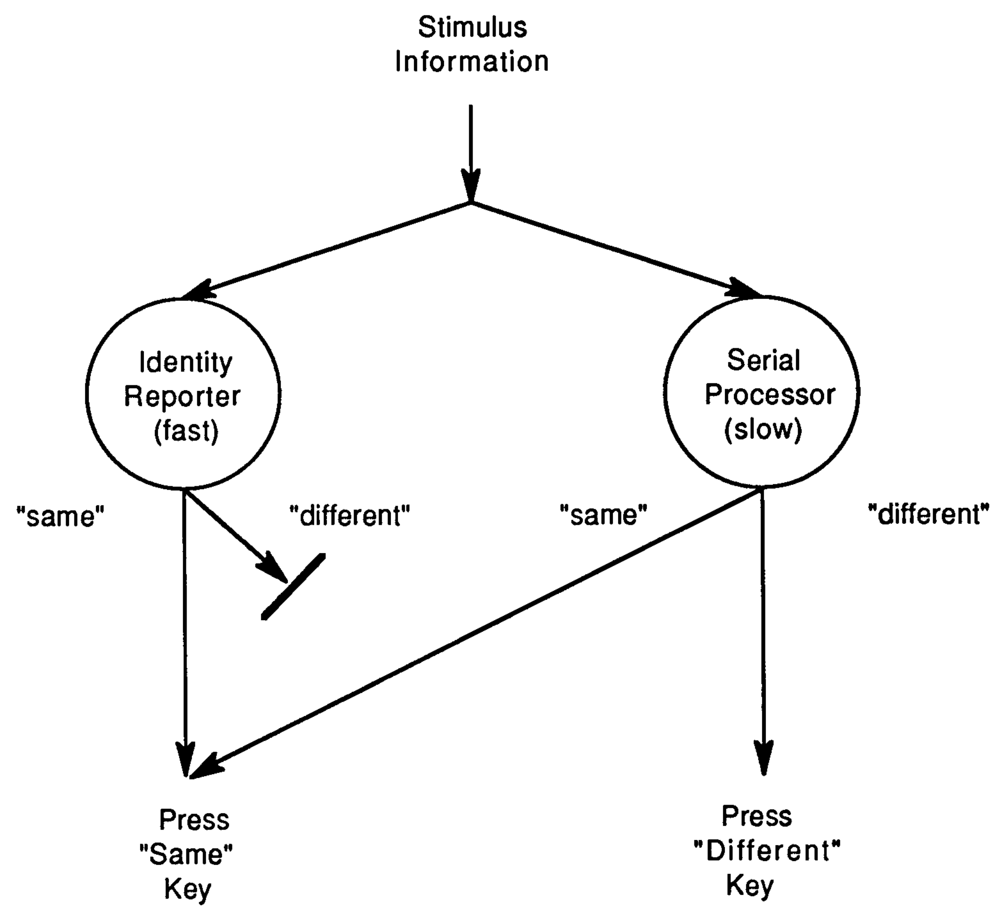
Bamber (1969) found that *different* responses to strings of between one to four consonants appeared to indicate a serial self–terminating search. However, RTs to *same* responses seemed incompatible with this model and he proposed a two–process model where the *identity reporter* process is assumed to be faster than the concurrent *serial processor*. This identity reporter process was proposed to explain the fast *same* advantage he observed. Subjects were attending faster to *same* stimuli than to *different* stimuli. This is contrary to expectation as comparison of multidimensional *same* stimuli would seem to involve comparison along a number of dimensions to ensure that the stimuli are not different, and it would seem reasonable to presume that correct *different* responses would be faster because fewer comparisons are required than for correct *same* responses. However, this advantage for *different* responses was not found by Bamber in his experiment.

Bamber's model assumes that processing is bidimensional and the *identity reporter* processor has only one *same* response. He argues therefore that the identity reporter is faster than the serial processor and that both processes occur concurrently. This complex processing model demonstrates that simple serial processing of the Sternberg model discussed above is not adequate to explain the data obtained by experimenters. A flow diagram of Bamber's model is produced in Figure 2.4.

Figure 2.4

<u>Flow diagram of the two–process model</u>

<u>for *same* and *different* responses</u>



(from Bamber, 1969; p. 172)

The diagram above (accurately reproduced from the original) is somewhat misleading in that the response times of both processes appear similar. The faster processing of the *identity reporter* might have been indicated with a shorter line in the diagram than the serial processor. It is assumed that the *identity reporter* process has only one output in this model, and that no output whatsoever if produced if the stimuli being compared are not the *same*. Bamber points out that this could lead to the objection that the non–emission of the *same* signal would be equivalent to the emission of the *different* response. Thus the fast *identity reporter* processor could initiate a *different* response before the serial processor. However, he argues that subjects must wait for the *same* signal before deciding that there is none. It is difficult to draw inferences from this study as the applicability of this processing model to music cognition has not formerly been investigated.

2.3.2. Factors affecting speed and accuracy

Krueger (1978) investigated the processing of multidimensional stimuli with single pattern (both geometric and the letters F, G, J, K, and L) and multiletter strings of not more than four letters (*e.g.* CXR and CDT). He confirmed the finding of earlier experiments that *same* judgments are more efficient in that they are generally made faster than *different* judgments (*cf.* Nickerson, 1972). This finding is inconsistent with the expectation that a *different* response can be made as soon as the different aspect is realised when two *different* stimuli are compared, and that the comparison of two *same* stimuli demands exhaustive processing in requiring the serial processing of all elements to confirm that there is no difference.

However, *different* judgments seem more efficient when accuracy is examined as subjects are more likely to respond *different* to a same comparison than to respond *same* to a different comparison. Krueger proposed his *noisy–operator theory* (Krueger, 1978) to explain longer RTs for *different* stimuli. The theory assumes that the comparison process involves a number of passes, during which features of the stimuli are either matched as identical or non–identical according to some criterion, which can be adjusted after each pass. The incorrect attribution of a feature as either matching or non–matching is considered as *noise* in the comparison process. Krueger postulates that internal noise is responsible for making a *same* comparison look different, and that the rechecking of *different* stimuli has been proposed to explain the RT advantage for *same* responses.

Ratcliff (1981) found that comparison of similar five–letter strings of consonants, where the comparison was a permutation of the original string, produced long RTs and low accuracy. Subjects were asked to respond *different* either (i) if one or more of the letters in the initial string were replaced in the comparison test string by new letters, or (ii) if two of the letters in the initial string were interchanged in the test string. He found that when adjacent letters were interchanged, RT was longer and accuracy lower than when non–adjacent letters were switched. From this he proposed that test strings are not compared letter by letter, but that the two are compared in memory to assess the amount of overlap between the two stimuli. From this it might be assumed that stimuli which are similar require more exhaustive processing.

Krueger's *noisy–operator* theory (Krueger, 1978) can explain the results of experiments which have attempted to bias responses to either a *cautious same* or *cautious different* response. For example, the speed and accuracy of responses to *same* and *different* letter strings of four consonants was manipulated by Ratcliff and Hacker (1981). Subjects were tested under two biased conditions. When subjects were instructed to respond *same* only when sure (*i.e. cautious same* condition), *same* judgments were slower than *different* judgments (*i.e.* the mean *same* response time was 573 milliseconds, whereas the mean *different* response time was 515 milliseconds). This relationship was reversed in the cautious *different* condition when subjects were instructed to respond *different* only when sure (*i.e.* the *cautious different* condition produced a mean *same* response time of 472 milliseconds, whereas the mean *different* response time was 582 milliseconds). This finding that RTs were sensitive to criterion manipulation is important, and Ratcliff and Hacker therefore argued that RTs as an absolute measure of processing should be interpreted with caution.

However, Procter and Rao (1982) were critical of the conclusions of this experiment by Ratcliff and Hacker (1981). Procter and Rao proposed that the RT differences were not attributable to bias factors and that the procedure was therefore appropriate for the examination of same–different processing. Procter and Rao point out that the differences between the mean RTs of *same* and *different* under each bias condition are distinct: 110 milliseconds separates the *same* and *different* RTs in the *cautious same* condition, whereas 58 seconds separates the *cautious* *different* condition response times for *same* and *different* stimuli.

Ratcliff and Hacker (1982), in reply to these criticisms of Proctor and Rao, point out that the two bias conditions are not equal as error rates between the two conditions are different. In the *cautious same* condition the probability of a false *same* response was .076 and false *different* .109, whereas in the *cautious different* condition the probability of a false *same* response was .139 and a false *different* .033. This variability of error rates shows that accuracy may also be subject to experimental manipulation. Ratcliff and Hacker reaffirmed that processing models should not place undue emphasis on differences between positive and negative responses as an absolute measure of processing since RTs in same–different tasks may be subject to experimental bias.

Experiments provide no doubt that RT experiments of letter recognition are subject to criterion manipulation and experimental effects. This possible manipulation in an important factor which must receive consideration in RT experimental design using musical materials.

2.3.3 Serial order effects

Proctor *et al.* (1991) have examined responses to *same* and *different* multiletter strings where the comparison string has the same letters as the original string but in different positions. They found that subjects asked to report *same* only when the order of letters was the same (the *order* task) produced left–to–right serial order effects, whereas subjects asked to report *same* regardless of the position of the letters produced U–shaped serial effects. This can be interpreted as suggesting processing from both ends of the stimulus when letter order is inconsequential.

Similar U–shaped profiles have been obtained in experiments with music materials, although not using RT as the dependent variable. Roberts (1986), for instance, has examined recall memory for melodic and harmonic music materials. The profiles for mean correct recall of eight–note melodies show the effects of both *primacy* (superior recall of initial items in the list) and *recency* (superior recall of terminal positions in the list of items). This memory superiority factor must be considered when devising musical experimental materials as the position within the musical sequence affects both memorability and consequently RT.

2.3.4 Visual domain processing

Proctor has also investigated RTs to pattern matching in the visual domain. Symmetry about the vertical axis is a characteristic of same pairs of patterns and this has been shown to facilitate the *same* response in RT experiments in the visual domain. Proctor *et al.* (1990) tested *same* and *different* matching of oval or racetrack patterns with backgrounds of non–parallel lines which were either symmetrical or asymmetrical. It was hypothesised that symmetric backgrounds would provide extraneous evidence for *same* responses and that asymmetric backgrounds would provide extraneous evidence for *different* responses. While their hypothesis was supported by blocks of trials in which all backgrounds were of the same type, they found that mixed random presentation affected the relative weighting of criteria adopted for *same* and *different* and induced the adoption of *compromise criteria* by subjects which takes account of the background noise.

Watanabe (1990) explored the same–different task in the visual domain with comparisons across the dimensions of form, size, orientation and colour. He used RT to examine the effect of irrelevant differences as a function of the relations between relevant and irrelevant dimensions. Although the visual domain is not directly comparable to the auditory domain, this experiment indicates that processing may occur along a number of dimensions even outside the delimited task itself, and that changes in background *noise* (*noisy–operator theory*, Krueger, 1978) might affect the RT rather than the relevant dimension itself. If the same–different response task is applied to musical stimuli then this must be considered as a potential experimental factor. If stimuli possess too many musical dimensions then this may adversely affect the analysis of RT. For example, a harmonic accompaniment to a melody might detract from the melodic characteristics being compared. Similarly, a rhythmic aspect of the music might influence the perception of the relative importance of certain notes of a melody. Palmer and Krumhansl (1987a, 1987b), for example, have found that pitch and temporal aspects of musical phrases are interdependent, and Brown (1985) has demonstrated that intervallic relationships in stimuli are subject to time order dependencies

2.3.5. Summary

Experiments ascertaining RTs to *same* or *different* responses to multiletter strings give insights into cognitive processes, but the perception and cognition of musical stimuli is not analogous to visual or linguistic processing. Non–semantic multiletter strings do not invoke higher order cognitive processes that may be invoked by word recognition. Such stimuli as non–word strings of letters do not necessarily instantiate a schema, *i.e.* a set of expectancies which inform future responses. Such strings of letters may not be contextualised within any higher–order cognitive structure. Thus RTs for visually presented letters are not directly comparable to musical stimuli which might instantiate a tonal schema, *i.e.* a tonal hierarchy. However, a word is a semantic unit in that it is representational and evokes a schema. What might constitute equivalent semantic units in music is not clear, although a melodic phrase might be regarded as a musical semantic unit if it instantiates a schema. However, this would be dependent on the length of the phrase.

RT experiments in the visual and linguistic domain provide only qualified answers to more general cognitive processes. More questions are posed by a consideration of the research literature than are answered. For instance, which models of RT best describe music processing? How does Bamber's *identity reporter* mechanism relate to music perception and is the faster *same* response advantage upheld by musical perception? If so, what is the nature of the *identity reporter* and does it relate to higher–order cognitive processing mechanism? Does Krueger's *noisy–operator* theory as a sequential–sampling model of cognition pertain to the comparison of music stimuli? Do RTs decrease with increasing age? Some of these questions are answered by the RT experiments using musical stimuli described in the following sections.

2.4. Chronometric studies of the musical interval sense

Balzano (1977) investigated the perceptual reality of chroma and scale–step using RT measures. His methodology was to visually present a musical interval name followed 1.6 seconds later by a musical interval stimulus. Subjects were required to declare whether the two stimuli were the *same* or *different* by pressing one of two keys. Some of his experiments presented melodic intervals (*i.e.* two successively presented tones) as he was interested in the context–generating effects of melodic presentation. He found that harmonic intervals *(i.e.* two simultaneously presented tones) generated more errors, and generally took more time to process than melodic intervals.

Subjects were all musically experienced adults (18 musically inclined listeners from Stanford University). RTs from this age group were mostly in the region of 500–900 milliseconds from the beginning of the sound to the keypress. Balzano applied a multidimensional analysis to the RTs and found that the two–dimensional configuration gave an approximation to the chroma circle (Revesz, 1954; Shepard, 1964). The configuration of the intervals resulted in a circular arrangement with the intervals contained within the octave appearing in ascending order (although not equidistant) from smaller to larger intervals. From this arrangement Balzano inferred that musical intervals are not unidimensional percepts varying only in width but have another cyclical component known as *chroma.* since he obtained a configuration approximating to the *chroma circle.*

From a methodological viewpoint, the *same* and *different* binary choice categorisation is associated with much previous research using RTs. However, one of the most significant problems in interpreting Balzano's findings relates to the non–contextual basis of some of the stimuli. Balzano's work predates the important contributions of Krumhansl's tonal hierarchy theory (Krumhansl, 1979) and Butler's theory of intervallic rivalry (Butler, 1989). Balzano's intervals are presented without any musical context and as such are not really helpful in revealing music cognition strategies with children. What is needed is a contextualised study of musical intervals insofar as they relate to music cognition.

The most important finding of Balzano is the perceptual salience of scale step interval. He examined the notion that semitone width is not constant but related to the scalic properties of the intervals. The visually presented label and musically presented interval were either classified as i) scale–step equivalent, or ii) not scale–step equivalent, related to the degrees of the scale forming the semitone interval. For example, although the interval of the major second (*i.e.* A–B) and the minor third (*i.e.* A–C) possess a semitone difference, as the upper notes (B and C) represent different degrees of the scale this comparison is termed not scale–step equivalent.[1] This type of comparison was contrasted with intervals which were classified as scale–step equivalent, such as the minor third (A–C) and the major third (A–C sharp). Balzano found a 256 millisecond difference ($p<0.001$) between the mean RTs of pairs of

[1] Balzano used A (440 Hz) as a base tone rendering his Experiments 1–3 in the key of A.

intervals that shared the same scale step (*i.e.* 1014 msec.) and pairs of intervals that were one semitone apart but not scale step equivalent (*i.e.* 758 msec.). In other words, it took longer to distinguish between intervals that shared the same scalic descriptor. This led to the conclusion that intervals that are scale–step equivalent are more similar than those that are not scale–step equivalent.

Balzano also investigated the discrimination of what he termed high and low level questions. A high level question demanded that subjects respond *same* to an interval if verbal and musical presentations were scale step equivalent (*e.g.* major second and minor second) whereas a low level question asked subjects to respond *same* only if the heard interval matched a specific intervallic description (*e.g.* a minor third). He found to his surprise that subjects were significantly quicker to respond *same* to paired visual and aural stimuli that shared the same scale step, the higher level task, than for intervals precisely specified in advance. For example, subjects were significantly faster ($F(1,9) = 8.14$, $p < 0.025$) to respond to a minor third when listening for either a minor or major third than when they were listening only for a minor third.

These findings led Balzano to argue strongly for the perceptual primacy of the scale step interval. In other words, his musically inclined listeners were aware that a particular interval was categorised as a second before they were able to classify it as either a major or minor second. This might be explained by listeners abstracting intervals to a scalar schema in the first instance before definitive classification can follow.

Balzano's results seem to indicate that the perception of musical intervals is dependent on abstraction to a scalar schema. It would be important to ascertain if children utilise the same processing strategy as adults for the experimental investigation proposed here relating to children's cognition of tonal organisation. If the scale step has perceptual primacy, a study of the acquisition of scalar schema might explain children's perception and cognition of pitch relationships. No researchers have yet used RT measures to explore children's acquisition of scalar schema.

2.5. Stage reduction theory and music cognition

Fiske has investigated cognition strategies in music listening using RTs. However, like Balzano, he has used experienced musicians as subjects and has not explored developmental factors or considered possible abstraction to a tonal schema.

Fiske (1982a, 1982b) applied chronometric analysis to the music listening process with music students. He adopted a binary–choice response (between *same* and *different*) for three progressively more difficult tasks. His first experimental condition was unidimensional in requiring the detection of a tonal discrepancy between pairs of isochronous tonal patterns or the detection of a rhythmic discrepancy between pairs of rhythmic patterns presented entirely on the same pitch of g' (*i.e.* task one). He considered this task easier than the bi–dimensional detection of either a tonal or rhythmic discrepancy between tonal–rhythmic patterns of seven non–repeated notes (*i.e.* task two). He considered that the identification of the type of discrepancy (either tonal or rhythmic) in seven–note melodies was the most difficult task of all (*i.e.* task three). Subjects were randomly assigned to one of the three tasks. He found statistically significant differences in RTs between the tasks.

In a development of this study, Fiske (1982b) used fragments of tonal melodies (usually about four bars long) expected to be familiar to the subjects. This contrasted with:

> *... the atonal, randomly generated phrases in the first experiment.*
>
> *(Fiske, 1982b, p. 33)*

This use of the word *atonal* in this context is problematic in that the tonal/rhythmic melody provided as an example is clearly in C major, using all seven notes of diatonic scale. What the word *atonal* means is not clear as these melodies are referred to as *diatonic* in the same review. In this second experiment, Fiske found that familiar non–complex items produced shorter RTs and lower error rates than the unfamiliar complex patterns. Although the more complex tasks required greater processing time, this second experiment with familiar short tonal melodies produced a different task hierarchy with shorter processing time for task one over task two in the first experiment being reversed in this second experiment.

One of the problems with experimentation of this kind is the construction of test materials which are equivalent perceptually, and Fiske acknowledges that perceptual non–equivalence might be partially responsible for differences obtained in RTs. Fiske's rule system for generating some of the test materials is arbitrary and it is not difficult to see why test materials may not be comparable. For example, the rule system for the construction of the items does not guarantee musically equivalent melodies, despite Fiske's postulation of perceptual equivalence:

An Electrocomp 101 synthesiser generating a flute-like timbre was employed to produce phrases consisting of seven non-repeated tones within the range of a major ninth with no interval greater than a fourth. Within these restrictions, the sequence of tones, both pitch-wise and melodically, was determined by a table of random numbers. Because the phrases were constructed from the same theoretical parameters, and since they were all sequenced by chance, all phrases were similar structurally and stylistically

(Fiske, 1982a, p. 38)

One aspect not investigated or discussed by Fiske is the complex relationship between the rhythmic and tonal structure of the melodies and the discrepancy between the two melodies. Furthermore, the effect of contour is ignored. The example he provides (1982a, *p.* 47) shows that the two tonal–rhythmic seven–note melodies have <u>different contours</u> generated by the changed direction of the altered note. This cannot lead to perceptual equivalence.

Fiske (1985) examined the formation and comparison of mental images in a series of three experiments. He was concerned to test whether mental comparison of musical images involved either:

(a) an active auditory–like image of P against which an incoming auditory pattern (P') is compared, or
(b) a set of 'instructions' used to test the agreement of the incoming pattern (P') with that of a recalled pattern (P).

(Fiske, 1985, p. 57)

This second process that Fiske outlines, dependent on the abstraction of some underlying features of the compared images, is consistent with abstraction to a schema, or generalised cognitive structure which selects and organises incoming information to a meaningful framework (Bartlett, 1932).

Fiske's first experiment of this study (Fiske, 1985) examined the effects of familiarity and he found that RTs were not significantly affected by this variable. Fiske's second experiment of this study examined contour, and he suggests that strategy (a) above, the image comparison strategy, explains his observed non–significant RT difference and significant differences in error rates. However, his third experiment seemed to support strategy (b) since there were both RT and error rate differences. Fiske was unable to draw firm conclusions from these experiments.

As noted above, different strategies have been proposed for *same* and *different* responses by Bamber (1969). It might be that either or both strategies that Fiske proposes are appropriate for certain conditions. If the comparison of two auditory patterns involves schema abstraction preceding the analysis along particular dimensions, as found for example by Balzano in the perception of musical intervals, then this might be comparable to some type of identity reporter mechanism as proposed by Bamber which produces a fast *same* response. This would relate directly to Fiske's strategy (b), the proposed *set of instructions* against which the incoming pattern is compared. If comparison at the schema level fails and rechecking becomes necessary then strategy (a) might be invoked and

Fiske's matching of an *auditory–like image* might be necessary to determine a *different* response.

Fiske's third experiment of this study (Fiske, 1985) found that more time was needed for responding to diatonic discrepancies than for chromatic ones, and that more errors occurred for diatonic discrepancies than chromatic. This finding would support the view that those comparisons that are not so easily abstracted to a tonal schema are more easy to detect as *different* at this stage. This is related to listening strategy (b), the set of instructions, involving higher level commonalities or underlying structure abstracted from the stimuli.

There seems to be some evidence to support a theory of a dual process for music listening. The processing of absolute interval information may be preceded by an holistic processing stage which examines the most salient features of the stimulus. A number of music psychologists, including Deutsch (1969), have examined the interaction of absolute interval information and contour information and proposed some kind of dual theory of music cognition. The identity reporter processing stage found in RT studies in the visual and linguistic domain might be applicable to music cognition. Fiske, however, interprets his RT studies in terms of Sternberg's simplistic stage reduction theory of separate component mental processes.

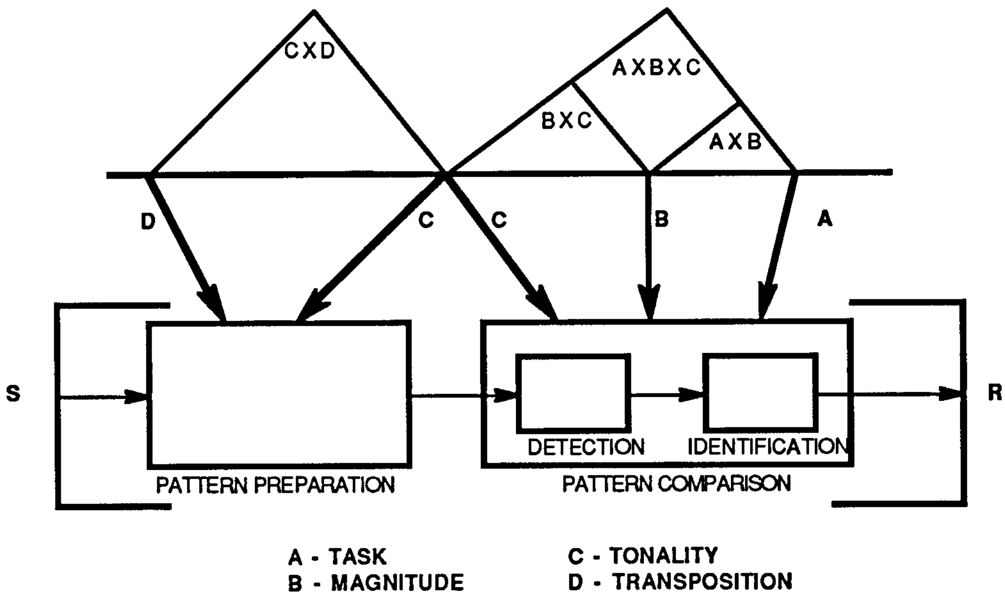
Fiske (1987) investigated the relationship between certain variables such as **task** (two levels – between *pattern detection discrepancy* and *pattern detection identification* of tonal or rhythmic discrepancies), **tonality** (two levels – between diatonic and chromatic), **interval magnitude** (two levels – seconds and sixths), and **transposition** (two levels – transposed to a fifth higher than the original and untransposed) in comparing ten–note melodies. Observing a number of significant differences, he formulated a complex model of music processing to explain his findings. He made the following propositions to explain the music listening process:

> *1. levels of task were found to interact with levels of discrepancy magnitude;*
> *2. levels of tonality were found to interact with levels of transposition;*
> *3. a three–way interaction was found between levels of task levels of discrepancy magnitude and levels of tonality.*
> *4. levels of discrepancy magnitude were found to interact with levels of tonality for the discrepancy detection task, but were not found to interact for the discrepancy identification task.*
>
> *(Fiske, 1990, p. 38)*

He formulated an elaborate pre–processing stage of pattern conciliation to embody these findings into some kind of processing model, although some of his results were inconsistent with some of his earlier experiments. His complex two–stage model (Figure 2.5) attempts to show the significant interactions he obtained (*e.g.* A×B, B×C, C×D, and A×B×C) between the four variables.

Figure 2.5

Pattern Comparison Model



(from Fiske, 1987, p. 36)

This model shows that tonality and transposition are more important in the pattern preparation stage than task and interval magnitude are in the comparison stage. Abstraction to a tonal schema as an initial process of comparison is supported by this model, as tonality and transposition are fundamental components of the pre–processing stage.

Fiske hopes that the model might serve as a basis for further empirical work. His research does not address the issue of the development of music cognition and listening comparison strategies which might be utilised by children since all his experiments used musically trained subjects (music undergraduates).

2.6. Reaction time and musical expectancy

2.6.1. Interval and contour

The relationship between interval and contour in the processing of melodies has been examined by Edworthy (1985a, 1985b) using RT responses. She used melodies with different numbers of equal–length notes (*i.e.* 3, 5, 7, 9, 11, 13, and 15 notes) which were all in C major. Her ten subjects (all experienced musicians with at least five years musical training) were required to detect changes in transposed versions to F sharp major. Edworthy found that melody length affected mean RT. Subjects were asked to undertake two tasks for each of the melody lengths. The *interval judgment* task required subjects to recognise if a pitch alteration (always to diatonic notes) occurred in the comparison melody. The *contour judgment* task required subjects to respond by pressing a button if they recognised a contour alteration in the comparison melody. For all melodies up to eleven notes long, mean RTs were shorter in the contour judgment than in the interval judgment. No significant difference between tasks was observed for 13–note melodies, and the 15–note melodies induced shorter RTs for *interval* than for *contour* judgment. From this Edworthy argued that contour information is immediately available on transposition but is increasing lost with melody length, whereas interval information is not as stable as contour information in the shorter melodies but is more resistant to forgetting in the longer melodies.

Edworthy considered the importance of the perceived tonal framework induced by stimuli. All of the transpositions of the comparison melodies were a tritone removed from the tonic of the standard melody. This would give the least degree of overlap of tonalities possible if the implied tonality of the comparison melodies was being compared to that of the standard melody. The accurate coding of the transposed comparison melody into constituent intervals seemed to improve with a clearly defined tonal context, although contour information was preserved. Edworthy proposed that interval information is lost, therefore, as the first few notes of the transposed comparison are perceived until a tonal context is generated by the stimulus. This means that the position of change in compared melodies is a critical factor.

Although Edworthy's study demonstrates the significant effect of melody length on processing strategy, the most important finding of her study is that the relative salience of interval and contour information is a function of the currently available tonal framework. However, none of Edworthy's melodies used the melodic three–note sequence of tritone plus one other context defining note which would unambiguously define a specific tonality. Her stimuli were centred around the notes of the tonic triad of either C major and F sharp major, depending on the implied tonality of the stimulus (*e.g.* CEDG/EFGD/CEFD/CGC^1 within the compass of the octave above middle C). Edworthy considers that until the key of the transposition is determined by the first few notes of the comparison melody, interval information is rendered imprecise. This seems to support the notion of a more complex processing strategy than a simple one stage comparison model where an incoming stimulus is matched with a conceptual template in a serial note–for–note fashion. Other global features of the musical stimulus (*e.g.* the suggested tonality or global contour) may form a pre–processing stage which might be followed by a more detailed analysis along specific dimensions.

2.6.2. Metric and harmonic rhythm

The studies of Smith and Cuddy (1989) looked at the relationship between metric and harmonic rhythm and the detection of pitch alterations in comparison melodic sequences. Like Edworthy's study, they utilised comparison transpositions a tritone removed (in F# major) from the standard (always in C major). They acknowledged that their experimental paradigm was adapted from Edworthy (1983) and many of the experimental sequences were taken from her study. An important feature of Smith and Cuddy's study is that sequences were learnt by multiple presentations (*e.g.* sequences were repeated 10 times in the first experiment) before the recognition trials. Moreover, pitch alterations were always within the key of the original sequence. It would have been more interesting for the present study if Smith and Cuddy had compared RT data for pitch alterations within the tonality of the standard melody with pitch alterations outside the tonality of the standard.

Smith and Cuddy (1989) are careful to document that they consulted both a teacher of music theory and a composer/theorist as independent musical judges. A decision to ensure equality of length of stimuli (*i.e.* 13 notes) between melodies in both 4/4 meter and 3/4 meter produced melodies in 3/4 meter which were 5 bars long! This might have affected the processing of stimuli and perhaps explain the 4/4 metrical structure superiority, although this is not commented on by the researchers. A psychologist might be happy to equate a four–bar 4/4 structure with a five–bar 3/4 structure since they look and sound equivalent in length. However, a musician would be aware that they are not perceptually equivalent. Figure 2.6 gives an example of some of the experimental materials used by Smith and Cuddy. The metric rhythm of either 3/4 or 4/4 was indicated by dynamic accents. The harmonic rhythm resulted from implied triadic changes instigated on the first beat of each bar.

Figure 2.6

Materials used by Smith and Cuddy (1989)

(from Smith and Cuddy, 1989, p. 460)

Smith and Cuddy found that responses were not always faster for those rhythms which were matched or for alterations placed on dynamic accents. They found that the metrical structure of 4/4 seemed to aid the abstraction of pitch content rather than the matching of the harmonic rhythm with that of the metrical structure. However, this may be a result of the perceptual non–equivalence of the stimuli. Subjects may have preferred the four–bar length of the 4/4 stimuli to the five-bar length of the 3/4 stimuli.

Assumptions concerning the musical characteristics of stimuli seem to be a recurrent problem in music psychology experiments. The problematic nature of classification of stimuli into *tonal* and *atonal* or *non–diatonic* was discussed in the previous chapter in relation to the tonal hierarchy (Krumhansl, 1979) and replicatory experiments (*e.g.* Cuddy and Badertscher, 1987) and has been noted in the work of Fiske (1982a, 1982b). There has been no shortage of experimenters adopting this simplistic distinction. Trehub's study of her so–called diatonic materials (Trehub *et al.*, 1986) betrays a misunderstanding of musical structure. The classification of stimuli by specified criteria often assumes some kind of perceptual equivalence and these inherent assumptions are not always made explicit by experimenters.

2.6.3. Priming

RT has also featured in experiments looking at the priming of chords. The paradigm used by Bharucha and Stoeckig (1986) involved subjects (Dartmouth College students) making a *true* or *false* response about a target chord which was either closely or distantly related to a previously heard chord. The previously heard chord acted as a *prime*, a stimulus that generates a set of expectancies. They found that major targets were identified significantly faster when related to the prime than when unrelated. One explanation of the finding that related chords take less time to process might be that abstraction to a tonal schema takes less time since the degree of overlap is greater between two tonalities which have more notes in common than unrelated tonalities which share few notes.

However the non–contextualised nature of the chords poses a problem. The overlapping higher harmonics of related chords might be partially responsible for their perceived relatedness. This would help to explain the unexpected result in that they found:

> *... no significant correlation between priming and musical training. This suggests that a decision task such as in–tune/out–of–tune can fruitfully tap the underlying processes of listeners of all levels of musical training.*
>
> *(Bharucha and Stoeckig, 1986, p. 410)*

A number of studies have found no apparent distinction between experienced and inexperienced music listeners in a variety of music tasks (*e.g.* Speer and Adams, 1985; Cross, Howell and West, 1985). The finding that musical training is not significant in a number of studies raises the question whether the processes being investigated actually operate perceptually in the cognition of music by experienced listeners. The processing of the particular task might not impinge on the processing mechanisms which would give an experienced musician task superiority in certain circumstances.

A further experiment by Bharucha and Stoeckig (1987) used RT to address the effect of overlapping frequency spectra in the processing of related chords. Their first experiment employed a priming paradigm where prime–target pairs shared no component notes and related pairs had overlapping frequency spectra. The follow–up experiment removed all overlapping frequency components. They found priming equally strong in both experiments and therefore proposed that since frequency–specific repetition priming does not account for harmonic expectation, a cognitive level of representation of spreading activation is supported. However, they again observed a lack of correlation between musical training and the size of the priming effect.

The connectionist approach to neural networks has been developed further by Bharucha (Bharucha, 1987, Bharucha and Olney, 1989). The ideas of networking and neural architecture are influences of computer system architecture and input–output models of cognitive representation. Bharucha's network representation of relationships between notes, chords and keys is little more than a reworking of the circle of fifths in a two–dimensional presentation. Bharucha considers whether his experiments support the notions of parallel or serial processing. He is keen to establish that serialism cannot explain the rapid response to some stimuli, but does not rule out the notion that certain processing stages might exemplify serial processing. In this respect, he supports the cascade model discussed at the beginning of this chapter.

Another priming study has examined the tonal hierarchy theory proposed by Krumhansl (1979) with response time measures. Janata and Reisberg (1988) employed a similar procedure to the Krumhansl study by employing either an ascending scale or tonic triad prime which was followed by a single note that subjects (all musically experienced adults) had to classify as belonging or not belonging to the suggested key. The definition of key was the major diatonic set. Janata and Reisberg measured the time taken to respond to the single note stimulus following the prime. The hypothesis was that notes:

> *that are more consonant or more stable within a given tonal context will be more quickly and more accurately recognised as "belonging" to that context.*
>
> *(Janata and Reisberg ,1988, p. 163)*

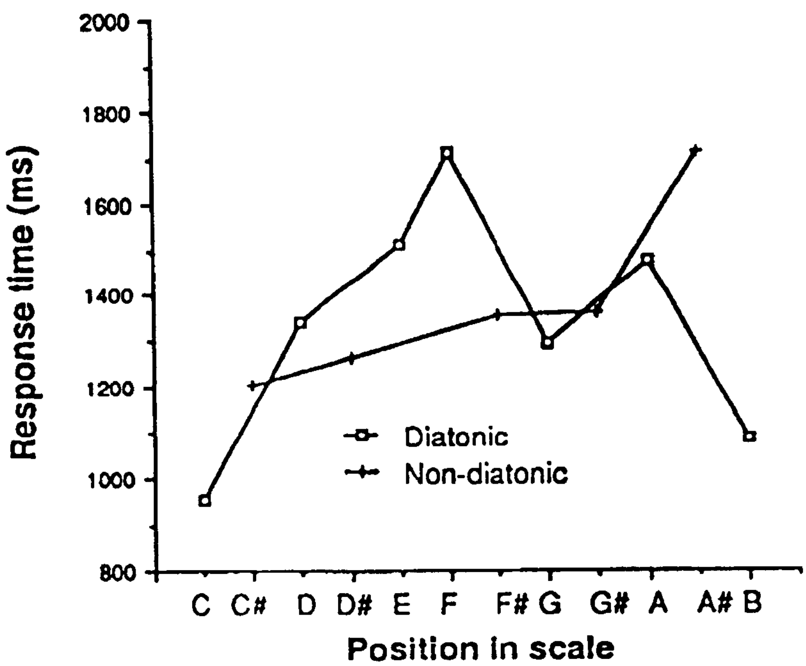
Trials using the scale prime presented each note of the ascending scale followed by the test tone, and trials using the chord prime presented the three notes of the tonic triad simultaneously. Surprisingly, the profiles they obtained for scale and tonic triad chord primes respectively were quite different. The response times for each position of the scale for each of the two conditions of scale and chord primes are shown in Figures 2.7 and 2.8. Responses to the leading note for the scale prime were comparatively short (just fractionally longer than for the tonic and faster than for any other note of the scale) and with fewer error responses than all other notes except the tonic (*i.e.* less than 10%). However, the leading note for the chord prime was responded to slowly (slower than all other notes except the subdominant which had a similar response time) and with more error responses than all other notes (*i.e.* more than 55%). Janata and Reisberg explain this observed difference in terms of a *recency* effect. The serially presented scalar stimulus here is obviously creating an effect as the leading note is more prominent in memory following the ascending scale prime. Janata and Reisberg might have used a descending scale context to control this, but they used only the ascending scale prime in the experiment.

This ascending prime has other consequences not fully explained by the investigators. A processing advantage for the mediant was observed in relation to the supertonic in the chord condition, although the scalar prime showed that the supertonic possessed an advantage over the mediant. They explain the shorter RT for the mediant over the supertonic in the chord condition as a priming effect of the mediant note within the tonic chord prime. The faster time observed for the supertonic over the mediant in the scale condition is explained as a *recency* effect linked to a self–terminating serial search procedure of the notes of the scalar prime.

One feature which was not commented upon in Janata and Reisberg's study was the comparatively longer RT for the subdominant in both the scale and chord conditions. The subdominant exhibited the longest response time in both the scale condition and the chord condition. The subdominant is proximal in key relation to the tonic and it is adjacent to it in the circle of fifths. These comparatively long RTs are difficult to explain as the subdominant chord is a stable chord within a given tonal context. The tonic triad prime could serve a dominant function in relation to the subdominant and therefore be considered as an expected resolution, particularly to the chord condition. Why this stable subdominant should give rise to more errors (about 45%) and greater response times than any other prime including the non–diatonic primes is difficult to understand as the data as provided (Figures 2.7 and 2.8) do not support Janata and Reisberg's hypothesis that notes that are more stable within a given tonal context are more quickly and more accurately recognised as "belonging" to that context.

Figure 2.7

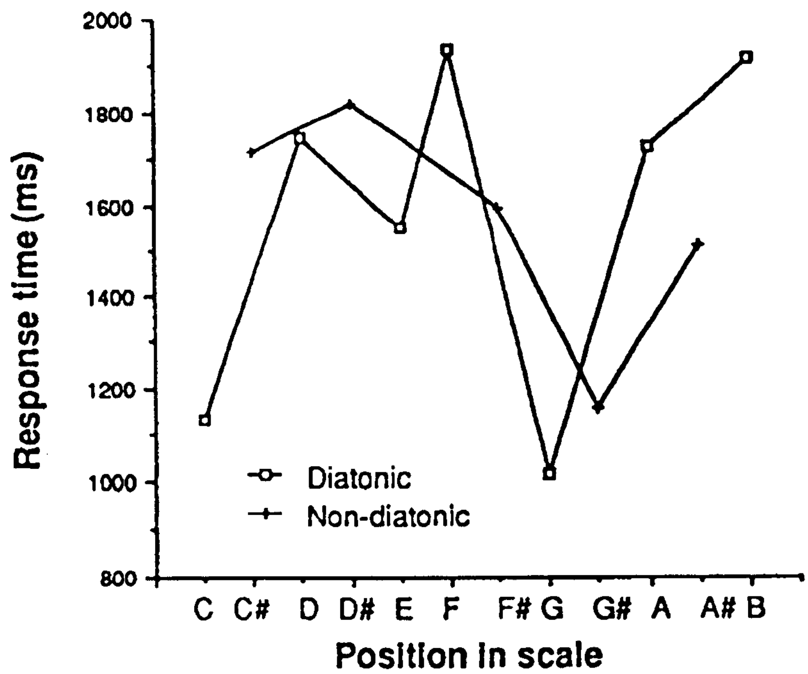
Response times for non–diatonic tones, SCALE condition, superimposed on the key profile for diatonic tones



(from Janata and Reisberg, 1988, p. 167)

Figure 2.8

Response times for non–diatonic tones, CHORD condition, superimposed on the key profile for diatonic tones



(from Janata and Reisberg, 1988, p. 167)

Janata and Reisberg's experiment assumes that the process of responding to a single note related to a prime invokes a schema which occurs during music listening. Such a schema may or may not be invoked by a stimulus, and the data is consistent with a two–stage music listening process in which the first stage attempts to process the information holistically before a second slower mechanism involving detailed individual feature comparisons becomes necessary to effect a comparison. If the first stage is abstraction to a tonal schema then those notes which are distant in the circle of fifths are likely to be recognised as different before a serial comparison search becomes necessary as a second stage. If abstraction is a two–stage process, then this might explain why some of the non–diatonic notes received shorter RTs than some of the diatonic notes in Janata and Reisberg's study. Such a two stage process might help to explain some of the unexpected results of the experiment and is consistent with RT models such as those of Bamber (1969) and the conflicting results of Fiske's experiments reported above.

Hubbard and Stoeckig (1988) have used a probe tone and priming paradigm with notes and chords to investigate subjects' ability to form a mental image of a chord or note one tone higher than the given cue. The mental image thus induced was then compared to a probe tone that was either the same as the image, harmonically closely related to the image, or harmonically distantly related to the image. They found that accuracy was greater for different *unrelated* targets then *related* targets. This would seem to support a processing model which involves an understanding of harmonic relations as in the circle of fifths. This notion of harmonic relatedness seems associated with RT, and Hubbard and Stoeckig explore the notion of distance along the circle of fifths, and argue that travelling equal distances along the circle of fifths yields approximately equivalent perceptual relatedness. They cite the results of Bharucha and Stoeckig (1987), discussed earlier in this chapter, as further evidence of the perceptual reality of harmonic relations.

2.7. Summary

The use of RTs in the experiments reported in this chapter has resulted in the formulation of a number of elaborate models to explain data obtained in both music and other domains.

Experimenters have found that neither the serial processing of the stage analysis model nor the parallel processing of the cascade model have been adequate to explain findings. For example, the cascade model has not been able to address the apparently complex relationship between speed and accuracy. Experimenters have found that subjects can choose to respond more quickly or slowly to stimuli at the expense of making more or fewer errors (Pachella, 1974). (Ratcliff, 1978, 1988) has proposed an alternative model, known as the *stochastic diffusion* model, which attempts to explain this relationship between greater speed and lesser accuracy. The model presumes that the initial response strength is set at a particular level according to some pre–determined criteria and that by adjusting this initial base level, an RT could be shortened because it would take less time to reach the response threshold. However, this drift towards a particular response threshold might result in prompting an error response since the *stochastic drift* (the random moving towards either of the two response thresholds) would have a greater chance of crossing the inappropriate threshold. This relationship between greater response times with low error rates and shorter response times with larger error rates is complex, and the simplistic discrete stage model of Sternberg does not always adequately explain differences in RTs.

More recent experimental approaches are suggesting a combination of priming and conventional RT responses to develop models which explain obtained RT data. For example, Meyer *et al.* (1988), have proposed a speed–accuracy decomposition technique whereby conventional RTs are combined with trials where subjects make prompted guesses before stimulus processing has been finished.

Many studies suggest that a one–process model is insufficient to explain findings. A number of models suggest some kind of holistic initial processing stage (*e.g.* Fiske's pattern preparation process) or, at least, some kind of faster mechanism which can precede serial searches along relevant dimensions (*e.g.* Bamber's identity reporter process). One of the most important questions which the present study seeks to address is whether children's music processing involves abstraction to a tonal schema as a component process of cognition.

The major problem with many of the experimental investigations of RTs to music stimuli is that non–contextualised stimuli are utilised and it is difficult to generalise from specific experiments, particularly when results seem inconsistent. There can be no doubt that chronometric analysis is a useful methodology for providing insight into cognitive processes, but assumptions concerning the processing model do need to be made explicit for the experiments to be interpreted. The differences in processing time observed in different experiments must be accounted for somehow, and significant differences in RT must be presumed to indicate some difference in cognitive processing. Although empirical work has examined a number of models, no clear conception of which models are best supported by musical stimuli has yet emerged. No experimentation with children has used chronometric measures to investigate the cognition of music. The following chapter considers methodological questions of using RT measures in relation to the principal models of musical pitch and examines how pitch perception might relate to abstraction to a tonal schema.

3. METHODOLOGICAL CONSIDERATIONS AND THE DEVELOPMENT OF EXPERIMENTAL MATERIALS

3.1. Processing Strategies

The elaborate models of cognitive processing derived from RT studies have shown that no simple serial model of processing yet proposed can fully explain the observed results. Many questions remain unclarified insofar as music cognition is concerned. For example, how are two sets of musical stimuli compared in memory? Do children compare two short melodies by holding the first in some short–term store and comparing the incoming comparison melody note–for–note? Is such serial processing, if it occurs, determined by mental rehearsal of the presented comparison? Is the first processing stage a pattern preparation stage as suggested by Fiske (1987) and outlined in the previous chapter?

An alternative processing strategy, particularly if the stimuli are too long to hold in short term store, might be that abstraction of certain characteristic features of the stimulus produces a matching or non–matching response generated by adaptation to a cognitive schema. An adaptive processing strategy of this type would relate incoming pitches to a cognitive reference point such as the tonic (Rosch, 1975; *cf.* chapter one). Edworthy (1983) found that longer sequences affected processing strategies. While note–for–note matching may well be employed for the comparison of easily remembered short sequences, more extended musical stimuli might promote processing on a more global level by encouraging reduction to a schematic representation of some kind. A comparison stimulus might contradict the schematic representation instantiated by a previous stimulus. The construction of appropriate experimental materials must consequently take account of stimuli length.

The effect of contour is another important consideration. One particular elaborate theory which has attempted to examine this is the *rule–recursion* theory of Boltz and Jones (1986). They examined the role of contour in variously structured melodies and considered the possibility that melodies are internally abstracted by observing structural similarities of their contour. While this may be true for melodies suggesting the same tonal centre, this may not be appropriate to describe comparison of melodies in contrasting tonalities. As the experimental materials used by Boltz and Jones were all taken from the C major diatonic set, the process of abstraction to a tonal schema was not considered. Moreover, the subjects were all sophisticated music listeners. Not surprisingly, they found no evidence to support their rather elaborate rule system, based exclusively on contour information. Clearly, both contour and pitch information are important to the cognition of melodies.

Stimuli that conform to well–learned patterns are processed on a more holistic level. For instance, Williams and Weisstein (1978) found that a pattern of lines in a visual display was easier to detect in a more complex visual pattern if the figure was meaningful. This effect has been called the *object superiority* effect. A similar effect has been found in word recognition. The *word–superiority* effect has been found for experienced readers. Young readers take as long to process non–word letter strings as they do to process real words, while experienced readers process real words faster (Juola, Schadler, Chabot, & McCaughey, 1978). The experimentation of Zenatti (1969), who used musical materials, appears consistent with the word–superiority effect. She found that children aged from about six to ten demonstrated superior discrimination of tonal sequences, as compared with atonal sequences. An examination of processing using RT with either different stimuli or differently aged children may demonstrate that abstraction to a tonal schema facilitates cognitive processing.

A number of researchers have explored schema abstraction using RT. Palmer (1977) has outlined a theoretical framework for perceptual representation using RT to look at the proposed internal representation of hierarchical networks. He required subjects to parse figures into their natural parts to discover if perceptual representations had a common structural organisation. He found that subjects exhibited a preference for certain configurations and quantified these subjective preferences as goodness ratings of parts within figures. He then used a *part–probe* methodology in which subjects had to recognise whether or not the segments of the part–probe were contained within the straight–line figure. He found significant differences in reaction time were generated by part–probes with different goodness ratings. Subjects identified more quickly those probes previously classified as having good parts within their figures. An adaptation of the methodology Palmer employed with straight–line figures using RT might fruitfully be extended and applied to musical stimuli.

Experimental work might seek to investigate cognitive representations of musical structure by using RT as an index of the internalisation of such structures. A chronometric study of music processing could explore the extent to which an hierarchical processing model would be appropriate to describe children's listening behaviour. Furthermore, if such a processing model is appropriate, it would be important to establish if there appear to be a developmental differences in music processing.

It is hypothesised that if subjects abstract musical pitches to an internalised tonal schema, then differential reaction times should result from the classification of stimuli which vary in their conformance to an established scalar schema. This proposition presupposes that abstraction to a tonal schema will take processing time.

A comparison of *different* stimuli whose conformance to the circle of fifths is similar could vary the position of the non–scalar note to examine whether the serial or holistic processing mechanisms, as discussed above, are perceptually salient. Such a strategy could also test for *primacy* and *recency* effects, particularly if both stimuli are equally well abstracted to a tonal schema. If the introduction of a further classifying element, *e.g.* contour, can be shown to shorten response time this is evidence for a hierarchy of perceptual relations. The notion of hierarchical structures conferring advantages in cognition and abstract representation has proved attractive in other psychological areas apart from music perception. Deutsch and Feroe's (1981) model is perhaps the most detailed attempt to explain and describe the internalisation and abstraction of music structures.

The role of the tonic is a particularly important consideration in stimulus generation for test materials. Experiments with younger subjects (Krumhansl & Keil, 1982; Trehub, 1987) have used tonally ambiguous short stimuli (*e.g.* a major triad in isolation suggesting a particular tonality by purporting to be the tonic triad). For instance, the triad of C major is tonally ambiguous in that it can represent the dominant of F major, the subdominant of G major, or the submediant of E minor, in addition to other diatonic functions in other keys. On the other hand, a chord such as the dominant seventh of G major (even as represented by the three notes D, F sharp and C) can most clearly be conceived as representing the tonality of G major. The expectancy frame that such a stimulus as a tonic triad produces is rather vague in that the mathematical possibility of certain notes appearing in the tonal schema is reduced, but not to a level of total certainty. However, the effect of priming by repeated context–generating materials is an important experimental effect which would influence the interpretation of a tonal centre. A mathematical model of the diatonic system has been proposed by

Agmon (1989), and attempts to represent the probabilities of certain notes occurring at any given time taking account of the complex nature of diatonic music.

The stimuli adopted for the initial experimentation utilised this ambiguity by attempting to examine the internalisation of pitch relations. Some understanding of the role of the tonic as a cognitive reference point was the primary concern of this initial experimentation. It was hypothesised that listeners could match successive pitches according to either an *exclusive* or an *inclusive* schema.

An *inclusive matching* schema would determine the relationships between pitches by comparison with a preconceived tonic. An *inclusive* theory would continually imply a mentally abstracted tonic to which all incoming pitches would be presented: incoming pitches would thus be *included* within the presumed tonal set. This processing strategy would be consistent with the tonal hierarchy theory.

Alternatively, an *exclusive rejection* schema would consider that the relationships between pitches would be not be determined by comparison with a preconceived tonic. An *exclusive* theory would not involve a mentally abstracted tonic to which all incoming pitches could be presented. An exclusive schema would eliminate certain prospective tonics by a process of rejection as pitches are presented, consistent with the presumed tonal strength of the stimulus. This processing strategy would be consistent with the intervallic rivalry theory.

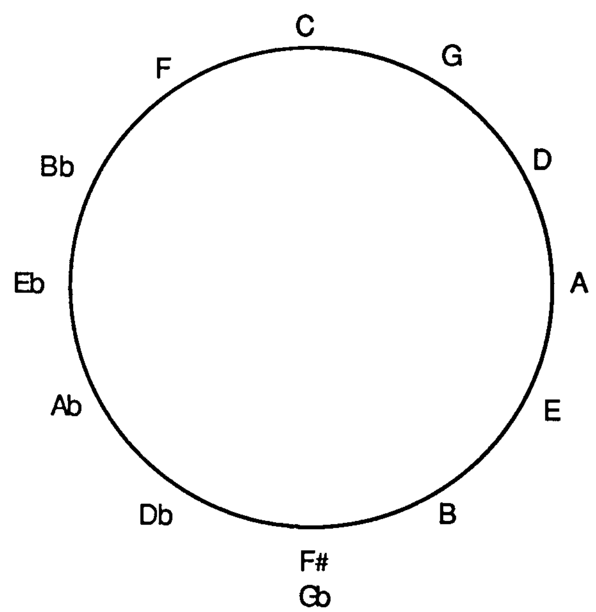
It is possible that these strategies are utilised concurrently dependent on the situation and tonal clarity of the presented pitches. For example, an ambiguous tonal stimulus might invoke an exclusive strategy rather than an inclusive one. Clarification of the notion of tonal ambiguity, particularly in how it relates to the development of experimental materials, is explored in the following section.

3.2. Pitch relationships

Different combinations of fixed pitches can imply different degrees of tonal ambiguity or strength. The tritone is the interval which most strongly characterises tonality in that it can belong to only two major scales. In other words, those scales whose tonics are a tritone apart share only two notes in common, e.g. the keys of C major and F sharp major have the notes F (enharmonic E sharp) and B in common. Tonal relations are shown clearly by the diagrammatic representation of the circle of fifths in Figure 3.1

Figure 3.1

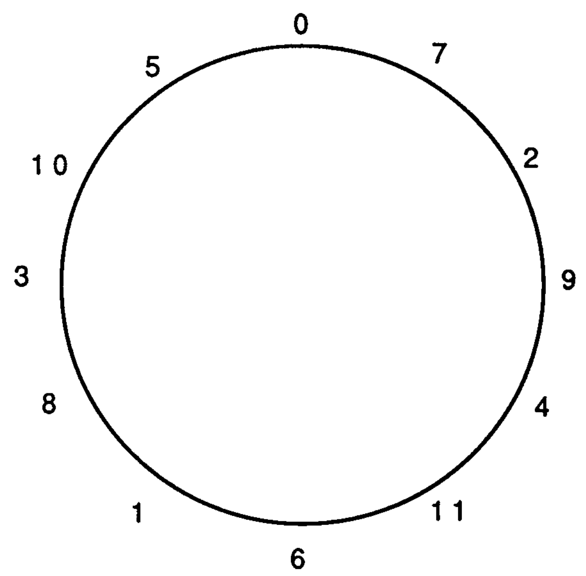
Diagrammatic representation of the circle of fifths



A representation where pitches are expressed numerically as notes of the chromatic scale (i.e. semitones) allows pitch relationships to be discussed without reference to specific pitches. This is shown in Figure 3.2.

Figure 3.2

Numerical representation of the circle of fifths



In this arrangement, keys are ordered by their relationships with each other in terms of the notes that they share. For instance, the scales of C and G share all their notes except one: if C has the notes 5,0,7,2,9,4,11 then G has the notes 0,7,2,9,4,11,6. Notes that are closer to each other represent tonalities that possess more notes in common. Notes which are proximal represent tonalities which have six notes in common, and notes which are directly opposite have tonalities which have the minimal relation of two possible notes in common. Those tonalities which are proximal are more related in music–theoretic description. Any seven adjacent notes around the edge of the circle form a major scale. Similarly, any five adjacent notes form the pentatonic scale which is characterised by its lack of semitones and less clear harmonic implications.

The tonal strength or *tonal specificity* of a musical stimulus can be determined by the spread of notes of the stimulus in relation to the circle of fifths. Each interval (and its inversion) is unique in specifying a particular number of tonalities. For example, the interval of a perfect fifth (or a perfect fourth by inversion) can occur in six sets of seven adjacent notes on the circle. For example, the notes 0 and 7 (represented numerically so as to avoid specific pitch names) occur in these scalic structures:

					0,	7,	2,	9,	4,	11,	6
				5,	0,	7,	2,	9,	4,	11	
			10,	5,	0,	7,	2,	9,	4		
		3,	10,	5,	0,	7,	2,	9			
	8,	3,	10,	5,	0,	7,	2				
1,	8,	3,	10,	5,	0,	7					

The interval of the major second (or minor seventh,by inversion) occurs within any five sets of seven notes. For instance, the notes 0 and 2 occur in the following groups:

				0,	7,	2,	9,	4,	11,	6
			5,	0,	7,	2,	9,	4,	11	
		10,	5,	0,	7,	2,	9,	4		
	3,	10,	5,	0,	7,	2,	9			
8,	3,	10,	5,	0,	7,	2				

The interval of the major sixth (or minor third, by inversion) occurs within four sets. For example, the notes 0 and 9 appear in:

			0,	7,	2,	9,	4,	11,	6
		5,	0,	7,	2,	9,	4,	11	
	10,	5,	0,	7,	2,	9,	4		
3,	10,	5,	0,	7,	2,	9			

The interval of the major third (or minor sixth, by inversion) occurs within three sets. For instance, the notes 0 and 4 occur in these sets:

		0,	7,	2,	9,	4,	11,	6
	5,	0,	7,	2,	9,	4,	11	
10,	5,	0,	7,	2,	9,	4		

The minor second (or major seventh, by inversion) appears within two sets only. For example, the notes 0 and 11 appear in these two sets:

	0,	7,	2,	9,	4,	11,	6
5,	0,	7,	2,	9,	4,	11	

Similarly, the tritone can occur only within two sets of notes but it suggests two completely different sets of notes. For instance the notes 0 and 6 appear in:

						0,	7,	2,	9,	4,	11,	6
6,	1,	8,	3,	10,	5,	0						

Although the tritone occurs in two sets of notes, the tritone and minor second do *not* have the same number of occurrences within any *one* tonal set. The tritone occurs once (*i.e.* between the subdominant and leading–note in one diatonic set) and the minor second twice (*i.e.* between the mediant and subdominant, and between the leading–note and tonic). The tritone is thus more tonally specific than the minor second. The *tonal specificity* of each interval can be ranked on seven different levels, in which zero is the strongest tonal indicator and six the weakest.

Uni.	Min. 2nd	Maj. 2nd	Min. 3rd	Maj. 3rd	Per. 4th	Aug. 4th	Per. 5th	Min. 6th	Maj. 6th	Min. 7th	Maj. 7th	Oct.
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
6	1	4	3	2	5	0	5	2	3	4	1	6

Such an arrangement is symmetrical. The tritone and minor second are more specific two–notes cues of tonality than other intervals.

Brown and Butler (1981) have investigated experimentally the importance of diatonic trichords as the minimum melodic cue–cell needed to unequivocally establish a particular tonality. They asked their musically experienced subjects to sing the tonic following a three–note stimulus. They found that two notes producing the interval of the augmented fourth with one other note was sufficient to fix a tonic in the listener's mind, even though the tonic may or may not have been heard in the stimulus. For example, they found that the tritone of D sharp and A sharp preceded by C sharp was sufficient to generate the implied tonic of E. They called this type of pitch string the *univalent cue–cell.* Other types of strings which had a number of possible interpretations were called *multivalent* strings.

Brown and Butler also found that temporal order of cue–cell components affected the sense of tonality implied by the stimulus, concluding that the structural hierarchy of the tonal set has perceptual validity only when it includes time–order dependencies. In other words, the perceptual salience of a set of notes is order–dependent and consequently relates to order of presentation. Palmer and Krumhansl (1987b) have also found that the perception of tonality is not independent of temporal considerations.

The demonstration of tonal relationships by Brown and Butler is important in that the tonic was abstracted from a stimulus which did not necessarily include the tonic. An investigation to discover whether children manifest the same sense of scalar conformance as adults is crucial to an understanding of the relationship of the circle of fifths to the perception of tonal melodies. The role of the tonic and its relation to the tritone seems crucial. The investigation must initially explore the notion that the the tonal specificity of a musical event or stimulus is determined by the spread of notes within the circle of fifths and that children (or adults, for that matter) may assimilate incoming information to a scalar schema influenced by this spread. A wider spread of notes around the circle may produce a more complete schema and greater tonal implication for that set.

Furthermore, the shortest stimulus from which a tonic can be extracted is an important consideration, particularly in the development of test materials. The task of analysing the effect of test materials is simplified if stimuli avoid redundancy and keep the information load to a minimum. There can be no doubt that a diatonic trichord which includes a tritone and one other note unequivocally identifies a particular tonality: this stimulus produces a stronger tonal image than the triadic or particularised scales that Krumhansl and her followers have used.

An investigation of the perception of pitch relations involves an examination of the cognitive structures which determine perceptual grouping. The most successful experimental methodology which might help to determine the rule systems which are used in cognition might explore the processing time taken for decisions concerning differences in similar stimuli. Diatonic trichords can be employed to generate stimuli with different tonal implications and varying degrees of tonal strength.

3.3. Cognitive Abstraction of Tonality

Initially, it was important to discover if children can abstract a tonic from a stimulus as short as a diatonic trichord. The Brown and Butler study (1981) asked subjects (all musically experienced) to sing the tonic or key–note following the presentation of three notes *(i.e.* the trichord) which were either *univalent* cue-cells (implying only one possible tonal centre) or *multivalent* strings (which implied two or more tonalities). The finding that the tonic can be abstracted from a stimulus which has not presented the tonic–note can support either the *inclusive matching theory* or *exclusive rejection theory,* postulated earlier in this chapter, dependent on the tonal strength or tonal specificity of the stimulus. A tonally unambiguous stimulus is likely to invoke an *inclusive matching* strategy whereas a tonally ambiguous stimulus is more likely to invoke an *exclusive rejection* schema.

This task of determining the implied tonic of a stimulus resembles the methodology of *pitch predominance.* This has been examined by Temko (1971), who required musically experienced subjects to sing the pitch they considered most important after the playing of a musical extract. This methodology, however, is inappropriate for younger children since they have a limited linguistic conceptual framework of the theoretical knowledge of tonality which would allow them to understand what is meant by the tonic or allow them to vocalise the tonic note. Children are unlikely to have reliably developed the necessary vocal–motor skills which would allow them to produce a vocal response. Clearly, production tasks of this kind are inappropriate for the experiments proposed here.

An investigation of tonality with children cannot rely on the ability of subjects to understand the complex linguistic framework required to comprehend what is meant by tonality. Children will not have the ability to relate such conceptual understanding to an explicit behaviour such as naming the implied tonic of a stimulus. The knowledge gained by children's interaction with the environment resulting in a developed schematic representation ready to interpret tonal music is largely implicit. The inability to describe this knowledge has little bearing on children's abilities to apply such knowledge to make sense of music. In language acquisition, structures are applied and developed subconsciously, and the lack of terminological apparatus needed to describe syntactical structures does not cause detriment to the intended expression. Similarly, schemata are developed in music largely by enculturation and interaction with existing schemata. A methodology which attempts to examine implicit knowledge is not therefore to be discounted simply because the terminology required to describe such musical features is unfamiliar to subjects. While explicit knowledge and description is useful for providing insights into how children and adults approach musical perception, it is inappropriate to attempt to extract such information directly from children, since different cognitive structures may be employed in the explanation process from those which make decisions regarding classification of musical stimuli.

Therefore, an experiment to see if children can abstract a tonic from a stimulus such as a trichord was an initial concern. The most tonally explicit diatonic trichord stimulus (embracing a tritone and one other contextual scale–defining note) might precede a presentation of probe–tones in an attempt to obtain rating profiles for the notes of the chromatic scale. Cuddy and Badertscher (1987) obtained tonality rating profiles from children aged between six to twelve years for each of the contexts of the major triad, the major ascending scale, and the diminished triad. They used an arpeggio of a diminished triad (B–D–F–B) which is the most explicit diatonic trichord since it contains a tritone and one other context–defining note. Cuddy and Badertscher recovered a flatter rating profile from children than the profile obtained from adults for probe–tones following the diminished triad. Children expressed no significant differences in preference for any notes. A repetition of the same experiment with university students yielded a more characteristic rating profile with preferences for the notes C, F sharp, and B. They conclude that:

> *The major–scale profile showed that adults were somewhat less influenced by pitch proximity than were the children. The cyclic properties of key structure were present in the adult major–scale profile. However, the scale was not as effective as the melodic major triad in recovering the tonal hierarchy. The diminished triad pattern did not recover a profile that was tonal.*
>
> *(Cuddy and Badertscher, 1987, p.618)*

This non-recovery of the tonal hierarchy with children is at odds with the expectation generated by the tonally specific tritone interval within the diminished triad stimulus. This would seem to indicate that intervallic rivalry is not the mechanism used by children in their music listening.

The rating profiles obtained by Cuddy and Badertscher in their experiment are different for each of the three context–defining stimuli. This suggests that the tonal hierarchy is not as stable as Krumhansl originally proposed, but that it is sensitive to the contextual properties of the stimulus. This would support the notion that the acoustical properties of the stimuli may affect the judgments of listeners. After all, the response which is required relates to a preference judgment according to the goodness of fit (supposedly musical) of the probe–tone in relation to the context. Such a judgment might be based, at least in part, on the acoustical properties of the notes concerned, the degree of *sensory consonance* (Terhardt, 1976, 1978) of the context–defining stimulus in relation to the probe–tone. The three notes of the major triad are much lower partials in the harmonic series than the three notes of the diminished triad. Therefore, according to accepted theory, the notes of the major triad produce a composite sound which is more concordant in physically producing fewer beats than the diminished triad. Preferential judgments based on such physical criteria are therefore not necessarily invoking cognitive structures which operate in music listening. Such a supposition might help to explain the differences in profiles obtained for different stimuli and the result that:

> *The essentially flat profile for the diminished triad suggested that this pattern conveyed no musical meaning for the children.*

(Cuddy & Badertscher, 1987, p. 616)

It is perhaps surprising that a triad or scale should have been found to invoke a stronger cognitive processing schema of tonal relations than the diminished triad (an unequivocal indicator of a specific tonality). It may be that the tonality profiles obtained by the probe–tone methodology have little to do with the cognitive structures employed in music listening. Cuddy and Badertscher's study used the diminished triad in root position which is not typical of the common musical usage of this chord: a first inversion orientation is more usual in musical contexts.

The usage of preference ratings, which rate the stability of the completion of stimuli on the seven point scale used by Krumhansl, is problematic with children. The concept of *'musical sense'* relates to rule systems by which notes are grouped to form larger coherent structures. These rule systems must be learned (consciously or subconsciously) by interaction with music, although predisposition to certain rule systems may be a possibility. An appropriate methodology for experiments with children should ideally ascertain information about such rule systems indirectly, rather than asking for an overt response involving preferential judgments.

If the scale, which is an explicit realisation of a particular tonality, is an important cognitive structural principle then it seems reasonable to propose that incoming pitch information is related to the set of pitches which could comprise a tonality, *i.e.* a scalar schema. If this is the case, then it should be possible to discover if listeners accommodate incoming pitches to an ever–changing tonal schema which excludes certain notes as a possible tonic as the stimulus progresses. This process would depend on the tonal ambiguity of the stimulus.

Children may respond more quickly in a discrimination task comparing stimuli that suggest different tonal sets than they would to stimuli that suggest the same tonal set, particularly if an holistic processing mechanism is perceptually salient. If the notes of two compared stimuli suggest the same tonality (and hence the same tonal schema or set of expected pitches), then an alternative or additional processing strategy may have to be employed to detect an alteration to the comparison. Such a task might involve pitch matching necessitating short–term memory storage of particular pitches if all presented pitches instantiate a particular tonal schema. However, the comparison of two pitches invoking different tonal schemata might be instantly recognised and remembered as different, although the precise pitches might not be able to be subsequently recalled.

The comparison of stimuli from the same tonal set may invoke cognitive structures which are more deeply nested hierarchically and which operate secondarily to the matching to a particular scale–set. For instance, as reported in the last chapter, Balzano (1977) proposed the perceptual *primacy* of the scale step in his research involving pitch matching of two stimuli. He demonstrated that incoming musical information is abstracted to the particular cognitive structure of the diatonic major scale. The expectancy frame of a stimulus is important in determining the processing strategy employed in subsequent comparison. Experimentation must attempt to invoke cognitive structures which operate in music listening.

A reasonable assumption might be that stimuli which have a more clearly defined tonality allow quicker judgments to be made if processing is hierarchical, and abstraction to a tonal schema precedes comparison of other features. This would assume that tonality processing is an initial processing strategy in a comparison procedure. An experimental method which presents two stimuli will allow a comparison to take place, with the first stimulus (*i.e.* the *standard*) defining a particular scalar schema and the second stimulus (*i.e.* the *comparison*) contradicting the previously established scalar schema. If the results show that it does in fact take less time to determine a difference, this may be evidence for global processing of an holistic nature. Analysis will not produce a tonal hierarchy rating profile of the type generated by Krumhansl, but the use of RT as a measure of the internalisation of stimuli will show if processing strategies are modified by the different tonal implications of stimuli.

A flow diagram of the cognitive model which it is proposed to evaluate is represented in Figure 3.3. The internalised representation of tonality is dependent on the tonal specificity of the stimulus. The cognitive representation resulting from an *exclusive rejection* strategy is likely to be different from that produced by an *inclusive matching* approach.

Figure 3.3

Flow diagram of proposed cognitive model of comparison process

---

STIMULUS

↓

COGNITIVE ABSTRACTION

↓

GENERATION OF TONAL SCHEMA

↓

MATCH TO EXISTING SCHEMA (ASSIMILATION)

↓

EXCLUSIVE REJECTION

OR

INCLUSIVE MATCHING

(DEPENDING ON THE *TONAL SPECIFICITY* OF THE STIMULUS)

↓

INTERNALISED REPRESENTATION OF TONALITY

↓

JUDGMENT RESPONSE OF SAME/DIFFERENT TONALITY

The experimental design which might test this model can be represented as in Figure 3.4.

Figure 3.4

Proposed experimental design of experiment

FIRST STIMULUS (*i.e.* Standard Stimulus)

CONTEXT ⇒ PROBE

(i.e. diatonic trichord)

⇓

SECOND STIMULUS (*i.e.* Comparison Stimulus)

CONTEXT ⇒ PROBE

(diatonic trichord of same/different tonality)

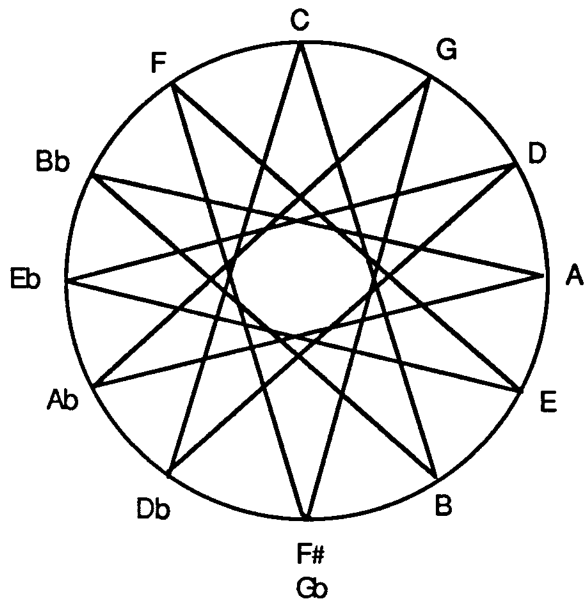
The use of a diatonic trichord cue-cell of the type used by Brown and Butler (1981) as a stimulus will produce an unequivocal tonal implication and make stimuli easier to classify tonally. That is, they should be easier to reject as *same* if they are from different scales and consequently suggest different tonal sets. For example, the note C, G and D which are proximal in the circle of fifths representation would prohibit an occurrence in the same scale of the note D flat (which is opposite in the circle of fifths).

It would be useful to attempt to obtain some type of value judgment (*i.e. which do you prefer?*) from subjects concerning preferences for stimuli. For example, subjects might express clearer preferences for those stimuli which come from the same schema rather than those which come from different tonalities. The idea of value judgment is an important musical consideration, but as it is likely to influence processing strategy and affect RT for responses it would be better to replay the stimuli a second time to ascertain this information. This might cause attention problems with younger subjects, so re–testing at a subsequent session might be considered.

Semitone discrimination has been a subject of investigation with a number of researchers (*e.g.* Trehub *et al.*, 1987). The experiment proposed for the study reported here allows an investigation of the semitone distance constituting the minor second interval, which is the next most tonally strong indicator of a particular tonality after the tritone. Semitones are maximally distant after the tritone in the circle of fifths representation as shown by the representation of Figure 3.5.

Figure 3.5

Semitone relations in the circle of fifths



In all of these diagrammatic representations of tonal implication as dictated by intervallic combinations, there is a danger of *a priori* invalid assumptions. However, the formulation of a clear conceptual basis from which an experiment can be conducted is necessary to test whether the theory of intervallic rivalry is perceptually salient for children.

3.4. Factors affecting development of experimental method

A number of concerns remained to be addressed before the test materials could be precisely formulated. Many of these were related to the differences between previous probe–tone experiments and the experiment proposed here.

The tonic note need not necessarily appear within the test stimulus at all with stimuli as short as diatonic trichords. However, some researchers have included the tonic in the test materials (*e.g.* the tonic is stated twice in the experiment by Cuddy and Badertscher; 1987). For instance, if the musical stimulus of a diminished triad (*e.g.* F sharp, A and C) is utilised this could be interpreted as G major although there is no G in the stimulus. If children hear tonally, at what point would they match these incoming pitches to an internalised reference point such as the tonic G? They would need to hear at least three notes to establish the tonality. However, the use of repeated trials with the same context–defining stimulus for each triad would allow *a priori* decision–making concerning the expected tonality of the stimuli. This idea implies the *inclusive* matching perceptual cognitive reference frame proposed earlier in this chapter.

An alternative strategy would presume that children match pitches to a perceptual hierarchy of the possibility of each note of the chromatic octave functioning as tonic. This cognitive strategy implies the idea of *exclusive* rejection whereby certain tonics are prohibited or less likely as the stimulus progresses.

The experiments reported in the next chapter tested the hypothesis that a tonic can be abstracted by children from such a stimulus as a diatonic trichord.

The effect of contour was another important consideration. The preservation of contour between comparison stimuli was desirable as contour change is an experimental variable which is best controlled. The length of stimuli was another factor which demanded attention. Edworthy (1985b) found that longer melodies were processed differently from shorter melodies. The minimum cue necessary to define a tonal context therefore seemed the logical starting point for stimuli generation. This meant that the stimuli need contain no more than three or four different notes. Moreover, temporal factors would have to be controlled in this experiment as this was another experimental variable which has been shown to affect results. For instance, Cross *et al.* (1983) found that manipulation of rhythm affected grouping, even with the same musical stimuli. The use of isochronous tones was considered not ideal in terms of musical context, but differences between stimuli comparisons due to rhythmic factors needed to be kept to a minimum.

The sense of key generated by groups of three or four different notes may be affected by the spread of those notes in relation to the circle of fifths. The use of cue–cells (which include a tritone and unambiguously define a particular tonality) and multivalent strings of pitches (which are ambiguous tonally) in a single experiment might have proved problematic. The number of items which need to be presented to compare all the different degrees of tonally defining contexts could be far too large for one experiment. The length of the experiment was an important consideration, particularly if younger children were to be employed as subjects. The task of matching pitches to the key defining context may be accomplished in a number of ways, particularly if the stimulus is short. The use of a distractor tone or pause between the two stimuli to be compared was considered as a possibility since such a technique would encourage listeners to memorise the stimulus. This in turn would induce a more processed and consequently deeper level of abstraction. However, the influence of the distractor tone on the RT was another factor which was considered undesirable and was therefore rejected.

The tonality of the test materials was considered. Krumhansl and others who have employed the probe–tone technique have maintained the same tonality throughout a set of test items. They have considered as problematic the assumed transpositional equivalence of stimuli. Since the stimuli for this experiment involve the abstraction of the tonic from the stimulus (which need not necessarily state the tonic), it seemed desirable to preserve the same tonality throughout the test items. After all, it was the relations between pitches that were being examined and not the pitches themselves. If subjects were to undertake the experiment individually using a computer, it would be possible to change the absolute pitches of the successive trials for a particular subject without changing the relative pitches of the notes within each trial. This would affect the tonal expectation between successive trials as each paired stimulus would imply a different tonal centre.

Many of the probe–tone tests rely on the experimental effect created by successive items reinforcing the same tonality. Jordan's (1987) demonstration of the assimilation of microtonal intervals to a tonal hierarchy might not have recovered the tonal hierarchy if successive test items had utilised different tonalities. It might be argued that we are observing an experimental effect here and not exclusively observing effects induced by the individual stimuli themselves. This can be addressed by the suggested methodology in that transposition of adjacent trials to different tonalities should negate this experimental effect. This important experimental effect needs to be acknowledged.

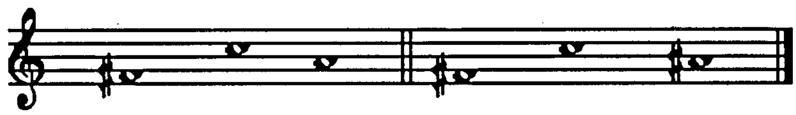
A number of response types were considered as appropriate to obtain evidence of cognitive processing from subjects. The preference ratings of the goodness of fit of the last note of a stimulus (*e.g.* as adopted by Krumhansl, 1979; and Cross *et al.*, 1983) give a good response scale for subsequent statistical analysis. However, it is somewhat inappropriate for the type of experiment proposed here as the adoption of the tritone, which is the most unstable interval in terms of tonal consonance, may affect the notion of what sounds 'right' melodically. The pictorial smiley–face differential used by Krumhansl and Keil (1982) is inappropriate for the same reasons. The use of RT with responses of *same* or *different* (possibly including a third category of *don't know*) seemed most appropriate. A further possibility might have been a variant of the matching technique from Personal Construct Theory in which the 'odd–man–out' (of three) can give insight into grouping if all permutations of possible groupings of stimuli are presented (*cf.* Ward, 1984).

While group testing of subjects was recognised as the easiest to organise, the nature of the required responses necessitated some kind of individual testing. The use of a computer–based testing environment had obvious advantages in that a precise chronometric measurement could be obtained, the testing environment could be identical for all subjects, and the computer could directly record subject responses thus minimising possible errors.

A pilot study using the cue–cell was considered suitable to assess the suitability of RT methodology with children. The *standard* (or initial) stimulus of each trial would comprise two notes generating a tonal context followed by a tone from the same tonal set. The *comparison* stimulus would consist of repetitions of the most tonally specific two–note context (*i.e.* a tritone) followed by a note which suggested either the *same* or a *different* tonal set. For example, the two three–note groups in Figure 3.6 suggests different tonal centres in that the last note of each three–note stimulus defines a different portion of the circle of fifths.

Figure 3.6

Musical example of three–note stimuli defining different portions of the circle of fifths



In this musical example, the first three–note stimulus suggests the tonality of G major whereas the second three–note stimulus suggests the tonality of D flat major (enharmonic C sharp major). It was hypothesised that stimuli suggesting different tonal sets would be discriminated as different more quickly than those from the same set, if children abstract pitches to a tonal schema before processing individual pitches or contour information. The experimental methodology sought to examine semitone and tone discrimination within tonal contexts (relating to the work of Trehub (1987) with infants discussed above).

The null hypothesis was as follows: *no differences will be observed in RTs of school–children to respond* ***same*** *or* ***different*** *to <u>standard</u> and <u>comparison</u> melodic stimuli which suggest the same tonality (i.e. are scale conformant) and those which suggest different tonalities.*

According to Brown and Butler (1981), the least ambiguous tonal cue is provided by the tritone and one other context defining note in the cue–cell. In terms of melodic contour, it seemed preferable on the basis of musical usage that the melodic tension established by the tritone should be resolved (*i.e.* the leading note should move to the tonic where possible).

For experimental purposes, one method of reducing the number of possibilities of note combinations would be to utilise the tritone as the context definition and the third note of the trichord as the changed comparison suffix. Each final note of the trichord could be linked to another stimulus which presented a note from the same tonality or another tonality, although the interval distance would be only a semitone, while preserving contour. It would be preferable to use only those trials which preserve similar contours.

It would be useful to present this test to subjects of different ages and levels of musical experience to examine the effect of age upon performance. Adults who are musically experienced, for example, could exhibit different listening strategies from children. The developmental nature of cognitive processes is of great interest. However, a pilot experiment with children of one age group contrasted with a group of adults would be more manageable.

Computer presentation of test items makes possible randomisation of order of items for each subject. This would negate any experimental effect arising from presentation order.

Although experimental stimuli finally adopted for the pilot studies could be criticised as atomistic or musically simplistic, their specific purpose was to effect a comparison of the effects of stimuli which might engage different tonal schemata. The information load would therefore kept to a minimum.

An examination of pitch predominance might have been useful in this test if, for example, a sung response had been obtained from subjects whereby they sing the tonic suggested by the stimuli (this task might be suitable for older subjects particularly). Both Temko (1971) and Brown and Butler (1981) have used this technique to good effect. However, this would only indicate which of the presented tones was prominent and not necessarily give an indication of the cognitive processes involved in the cognition of tonality. For example, as the implied tonic note is missing from each of the three–note stimuli in Figure 6.6, a sung response might not invoke a tonic response from subjects. Perhaps confusion might arise more frequently in the comparison of those stimuli whose final notes emanate from different tonalities.

The probe–tone methodology developed by Krumhansl has subsequently been developed by other experimenters. Clearly, if the results from the pilot experiment are at variance with those previously demonstrated this might prompt re–analysis in terms of a more–embracing theory of scalar relationships in some kind of expectancy model. Such an analysis might help to explain some of the inconsistencies noted by Speer and Adams (1985) and Speer and Meeks (1985) as well as throw light on the order of the developmental acquisition sequence which Krumhansl and Keil (1982) note as different from that proposed by Dowling (1982).

PILOT STUDIES: *SAME* OR *DIFFERENT* RESPONSES TO TRICHORDAL NOTE-STRINGS IN DIATONIC CONTEXT

4.1. Rationale of pilot study one

The experimental design of the pilot study comprised the randomised presentation of a number of *different* stimuli interspersed with a number of other stimuli which were the *same*. Half the trials suggested the *same* tonality and the other half suggested *different* tonalities. Equal numbers of *same* and *different* stimuli avoided the bias which could result from unequal numbers of *same* and *different* stimuli. In the design adopted for the pilot experiments, the initial standard stimulus of successive trials was always consistent with one particular tonal set. An alternative design could have included non-diatonic notes in the standard stimulus, but the cumulative effect of the reinforcement of the same tonality was the experimental effect which was being investigated. The tonal context was generated successively by the test items which served to establish the feeling of a consistent tonal centre, functioning in a similar fashion to the stimuli utilised by probe–tone methodology.

The first pilot study investigated children's discrimination of semitone change within paired melodic stimuli suggesting different tonal sets. The initial experimentation was undertaken with a limited number of trials.

Given the time-consuming nature of the testing (supervised individual responses), twelve trials were judged to provide a sufficient appraisal of responses from specified age groups and allowed the development of appropriate methodology.

A group of 8 to 9 year olds was chosen since by this age children should should be developing tonal schemata. The appropriateness of this age range is suggested by other research in tonality acquisition. For example, Imberty (1969) investigated the acquisition of tonality with children of various ages. He found that 6 year olds had little grasp of the import of cadences in harmonic structure, but noticed a significant improvement by the age of eight in the children's estimation of the *incompleteness* of a phrase without a cadence. Teplov (1966) also found that 8 year olds could discriminate between those melodies which were complete and those which did not end with a stable note such as the tonic.

Before more extensive experimental work could be undertaken, it was necessary to evaluate the perceptual salience of the interval of a semitone in this paired comparisons presentation. RT measurement could provide an indication of whether different processing strategies were being employed for pitch matching in the different contexts of *same* and *different*. The RT was measured from the start of the changed note in the second stimulus.

A further research question concerned the nature of any observed differences in RT. Stimuli from two contrasting tonal schemata could be recognised either more quickly or more slowly as *different*. This would be dependent on processing strategy, and the quantification of the nature of this difference would allow the formulation of further research questions.

The pilot studies, therefore, were concerned with:

i) the discrimination of semitone change,

ii) the relationship between *same* and *different* RT responses, and

iii) the tonal implication of the diatonic tritone and observed RT.

The experiment was designed to test the following null hypotheses:

1. Children will show no significant difference in correct/incorrect classification of semitone change within pairs of diatonic trichords, with each trichord incorporating the tritone interval.

2. Children will not exhibit significant differences in RT responses between paired stimuli which are different and those which are the same.

3. No significant differences in RT responses will be observed for those stimuli which comprise notes from differing tonalities.

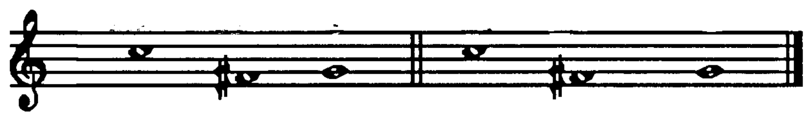
4.2. Method

4.2.1. Design:

The independent variable in the experiment was the nature of the stimulus under the experimental conditions. Paired stimuli, comprising two standard tones followed by one suffix tone, were presented in one of three conditions.

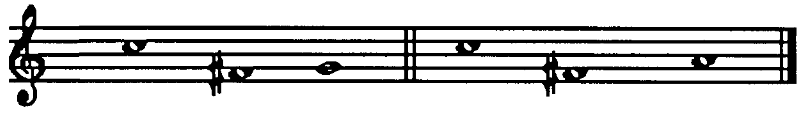
i) Condition 1: **same** suffix notes (six trials). In Trial 1, for example, the suffix notes of both the standard stimulus and comparison stimulus are identical. All **same** condition stimuli suggest G major.[1]

Trial 1



ii) Condition 2: **different** suffix notes but each trichord suggested the **same tonal set** (three trials). In Trial 7, for example, the suffix notes of both the standard stimulus (*i.e.* G) and the comparison stimulus (*i.e.* A) are consistent with G major.

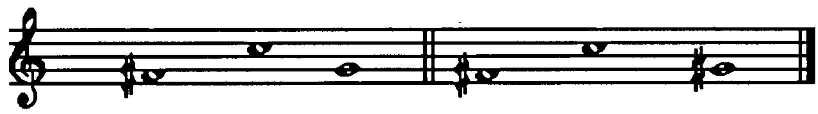
Trial 7



[1] Appendix II provides a musical representation of the experimental materials for all twelve trials.

iii) <u>Condition 3</u>: **different** suffix notes but each trichord suggested a **different tonal set** (three trials). In Trial 10, for example, the suffix note of the standard stimulus suggests G major, whereas the suffix note of the comparison stimulus (G sharp; enharmonic A flat) is consistent with D flat major (as G flat, A flat, and C).

Trial 10



The two dependent variables were i) the nature of the response (*i.e. same* or *different*) and ii) the time elapsing between onset of the changed note and the appropriate response. The response time was measured in hundredths of a second from the beginning of the changed note in the second comparison stimulus. Since the measurement of response times was in hundredths of a second, all response times are reported in centi-seconds.

The experimental design was repeated measures since the same subjects served under all conditions of the experiment. The order of presentation of test items was randomised by the computer to produce a different order for each subject, negating any experimental effect resulting from test familiarity and practice.

4.2.2. Subjects

The experiment was undertaken with 24 subjects (14 girls and 10 boys) aged between 8 and 9.[2] All subjects had normal hearing apart from one subject with considerable hearing loss who had recently had an operation to improve her hearing.

4.2.3. Apparatus and materials

The test was presented using a BBC microcomputer running a BASIC program which prompted the subjects for their responses and recorded the responses directly to disk for subsequent analysis.[3]

[2] The subjects constituted the second year junior class of an inner city Church of England School in Derby, England.

[3] The computer program which presented the trials is given in Appendix III.

4.2.4. Procedure

The test was administered to subjects individually in a room away from the distractions of the class. The verbal instructions given to subjects were as follows:

> "The computer will play two tunes, each with three notes, which will either be the *same* or *different*. Keep a finger of one of your hands between the **S** and **D** keys, and press either **S** (*for same*) or **D** (*for different*) on the computer keyboard as soon as you are sure."

The **S** and **D** keys are adjacent on the computer keyboard and the experimenter pointed to the appropriate keys on the computer keyboard.

In order to avoid possible reduction of experimental effect through test fatigue, the children did not undertake a practice test. However, the randomised presentation of test items for each subject would also help reduce any consequent experimental effect.

4.3. Results

4.3.1. Levels of performance

The results of the girl who had recently undergone an operation to improve her hearing were included in the analysis as she made no errors and her response times were close to the mean scores for each test item.

The number of correct and error responses to the various test items together with the observed proportion of the correct responses and the two-tailed binomial probability of these proportions occurring by chance (*i.e.* $p=q=0.5$) is presented in Table 4.1.

Table 4.1

Number of CORRECT and INCORRECT responses to the trials

Trial	Correct Responses	Error Responses	Observed proportion	Binomial *p*	
1	18	6	.75	.0227	*
2	20	4	.83	.0015	**
3	19	5	.79	.0066	**
4	21	3	.88	.0003	**
5	21	3	.88	.0003	**
6	20	4	.83	.0015	**
7	20	4	.83	.0015	**
8	20	4	.83	.0015	**
9	17	7	.71	.0639	
10	20	4	.83	.0015	**
11	19	5	.79	.0066	**
12	22	2	.92	.0000	**

N=24 (14 girls, 10 boys)

* $p<0.05$
** $p<0.01$

No significant difference was observed for error rates either between *same* and *different* stimuli (Sign Test, T=17, L=8, not significant) or between stimuli suggesting *same* or *different* tonalities (Sign Test, T=9, L=3, not significant). A comparison of error rates for the first six presented test items to each subject with the last six presented to each subject showed no significant difference (first six items=26 errors; last six items=23 errors). The error rate for the complete test was 17% (*i.e.* 49 errors out of 288 responses).

4.3.2. Reaction times

The first two items undertaken by each subject in each test were considered as practice items and removed from the analysis of RTs. As the trial order was different for each subject, practice items were randomly distributed across all trials. This reduced the influence of a possible training effect which might have resulted in faster reactions as the test progressed, although this effect is counterbalanced to a certain extent by the different randomised presentation of test items for each subject. Furthermore, error responses were also removed from the analysis.

The mean RT scores and standard deviations for correct responses to the trials under the three experimental conditions are shown in Table 4.2.[4]

[4] Each RT in centi–seconds is reported in rounded form to one decimal place since this gives an approximation to milliseconds. However, the computer programs which were specifically written to compute the means and standard deviations reported in the following pages carried numbers to many decimal places. As nine significant figures were always reported by the computer, between five and six decimal places were usually reported and carried for each value. This avoided the systematic error which would have been generated by rounding to the nearest integer or designated decimal places.

Table 4.2

Mean RTs and SD for CORRECT Responses[5]

	Condition	Trial	Mean RT (centi-seconds)	Standard Deviation (centi-seconds)
1.	**Same**	1	282.9	101.8
		2	279.2	40.1
		3	283.6	55.6
		4	314.3	111.6
		5	300.9	101.7
		6	304.4	114.6
2.	**Different**	7	260.3	25.6
	(Same	8	280.4	68.3
	Tonality)	9	290.3	67.3
3.	**Different**	10	266.5	37.1
	(Different	11	269.2	56.0
	Tonality)	12	257.9	27.4

The variability of mean response times demonstrated by the variance for each subject's CORRECT RT performance under each condition was examined by variance–ratio tests. Considerable significant differences in variance were observed between all three conditions. The variance of RT responses for condition 1 (same stimuli) and condition 2 (different stimuli suggesting the same tonality) were significantly different ($F(23,21)=4.2$, $p<0.01$). Responses in condition 1 (same stimuli) and condition 3 (different stimuli suggesting different tonality) were also significantly different ($F(23,21)=10.08$, $p<0.001$). The responses in condition 2 (different stimuli suggesting the same tonality) and condition 3 (different

[5] Trial numbers are for identification purposes only (see Appendix II) and do not refer to presentation order as trial order was randomised for each subject.

stimuli suggesting different tonality) were also significantly different ($F(21,21)=2.4$, $p<0.05$)

The mean RTs for the trials for each condition were analysed by a nonparametric Kruskal-Wallis ANOVA, considered appropriate since the sample size of each group was small and there were clear differences in variance between the conditions. The Kruskal-Wallis ANOVA confirmed that there were significant differences between the conditions ($\chi^2=6.064$, $p=0.048$).

The mean RTs and standard deviations for the three experimental conditions are shown in Table 4.3.

Table 4.3

Mean RTs and SD for the three experimental conditions

Condition		Mean	Standard Deviation
1.	**Same**	308.6	85.9
2.	**Different (Same Tonality)**	278.0	41.9
3.	**Different (Different Tonality)**	264.6	27.1

The means for each condition revealed that children showed an increasing facility to respond as the stimuli become distinct in tonal implication. Furthermore, the smaller standard deviations show that responses are less variable for tonally distinct comparisons.

4.4. Discussion

The unequal proportions of correct/incorrect *same* and *different* responses for the test items were significant (Table 4.1) for all of the trials except number 9 (which approached significance, *i.e.* p=0.06). The first null hypothesis was rejected as there were significant differences in the numbers of correct and incorrect responses for the paired melodic stimuli. This confirms that subjects were able to discriminate the interval of the semitone within the context of trichordal note-strings incorporating the tritone interval. Their choices were well above those expected by chance on an equiprobable classification task.

The mean response times and standard deviations show considerable differences between the various test conditions. The second null hypothesis was rejected as the Kruskal–Wallis ANOVA confirmed significant differences in the response times for the different conditions of *same* and *different*. The third null hypothesis was also rejected as the variance–ratio tests confirmed differences in variability between those different stimuli which suggested different tonalities. Subjects responded more quickly and with smaller variability to paired stimuli that suggested different tonalities. This is compatible with the hypothesis that stimuli are initially referred to a tonal schema before absolute or individual pitches are processed, and that classification of *same/different* can precede such absolute pitch matching if different schemata are suggested. The longer RTs for *same* conditions implies sequential hypothesis–checking for both condition 1 (the same stimuli) and condition 2 (different stimuli suggesting the same tonality). This would explain why it took longer to classify correctly condition 2 (two different stimuli suggesting the same tonal schema) than condition 3 (two different stimuli suggesting two different tonalities). Furthermore, this would explain why subjects demonstrated more variability in their scores for condition 1 and 2.

Most interestingly, the mean response times for condition 1 (the same stimuli) were longer than other conditions with considerable differences exhibited between subjects. The classification probably requires mental rehearsal of the internalised representation of the stimuli to discover any such changes. The fast *same* advantage, as hypothesised by Bamber (1969; *cf.* chapter two), to explain faster RTs to *same* stimuli, is not supported by the data obtained in this pilot study. The fact that *same* responses took much longer can be explained by rechecking the stimuli across a number of dimensions. The data suggest exhaustive processing for *same* stimuli.

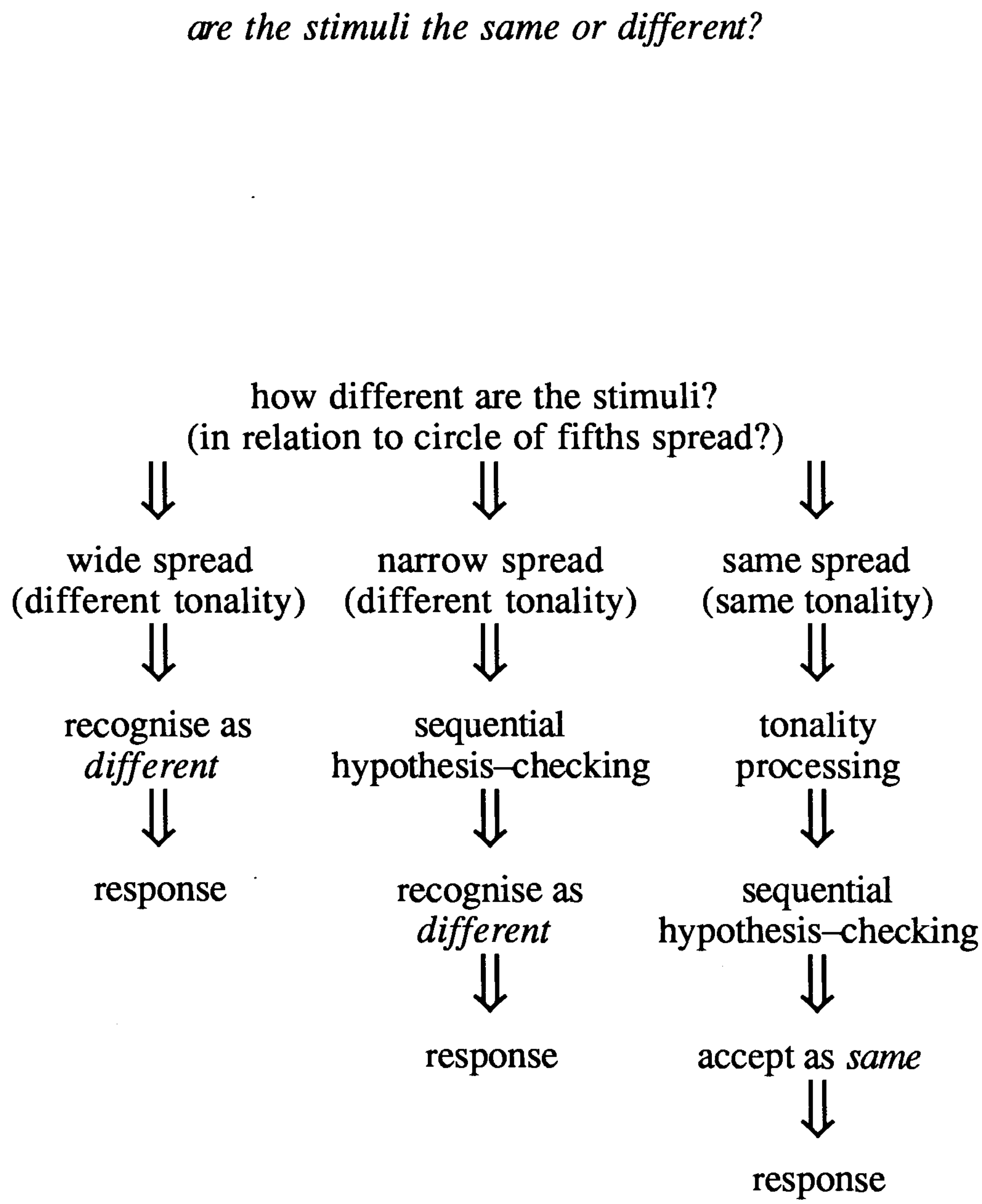
Significant differences in RT observed for different types of stimuli support the notion that different processing strategies are employed for pitch matching in different contexts. The fact that it takes comparatively less time to respond correctly to two *different* stimuli suggests that *same* processing involves comparisons along a number of dimensions. One possible interpretation of this difference is that a global abstraction to an internalised set of tonal relations precedes processing of absolute pitches. This supports the two stage abstraction process described in chapter three where abstraction to a tonal schema precedes processing of absolute pitches. Such an interpretation is consistent with the two–stage model proposed by Fiske (1987), which incorporates a pattern conciliation stage as a component process.

These results demonstrate that a diatonic trichord stimulus which includes a tritone is sufficient to generate a tonal context. This conflicts with the conclusions of of Cuddy and Badertscher (1987) who found that a diminished triad (which the tritone delimits) was insufficient to recover a profile from children that was tonal.

The results of this experiment leads to rejection of all null hypotheses and provides evidence for hierarchical processing. Those decisions which are made more quickly are hypothesised as operating at a different level within the hierarchy. A diagrammatic representation of the model of hierarchical processing suggested by this experiment might be represented as in Figure 4.1. The vertical arrows (⇓) represent processing time.

Figure 4.1

Diagrammatic representation of the model of hierarchical processing



A number of important questions are raised by this hypothesised processing model, not least the way in which this hierarchy might change developmentally as children mature (or at least their processing strategies might appear to change). For the purpose of comparison, the same test was administered to a group of adults. This was designed to investigate whether the same hierarchy of perceptual relationships would be manifested by adults and to what extent their processing strategies might differ from those of children.

4.5. Rationale of pilot study two

The second pilot study investigated the RTs of young adults to the same melodic stimuli incorporating semitone change and suggesting different tonalities employed in the first pilot study. The test was administered to 10 undergraduate students undertaking degree courses in music.[6] Subjects followed the same procedure adopted for children except that they undertook a practice test of five items prior to the administration of the actual test since it was considered they would not be affected by test fatigue. This meant that there was no need for them to be familiarised with the sounds of the test and they did not observe or listen to other subjects undertaking the test.

[6] Students were music undergraduates at a College of Higher Education in Derbyshire.

4.6. Results

4.6.1. Levels of performance

Musically experienced adults made many fewer errors than the children. Only two subjects produced error responses and errors included item 3 (both subjects), item 4 (one subject), and item 6 (one subject). No mistakes were induced by stimuli which were *different*. Stimuli which were the *same* produced four errors. The error rate for the complete test was 2.8% (*i.e.* 4 errors out of 144 responses).

4.6.2. Reaction times

The mean RT scores and standard deviations for correct responses to the test items under the three experimental conditions are shown in Table 4.4.

Table 4.4

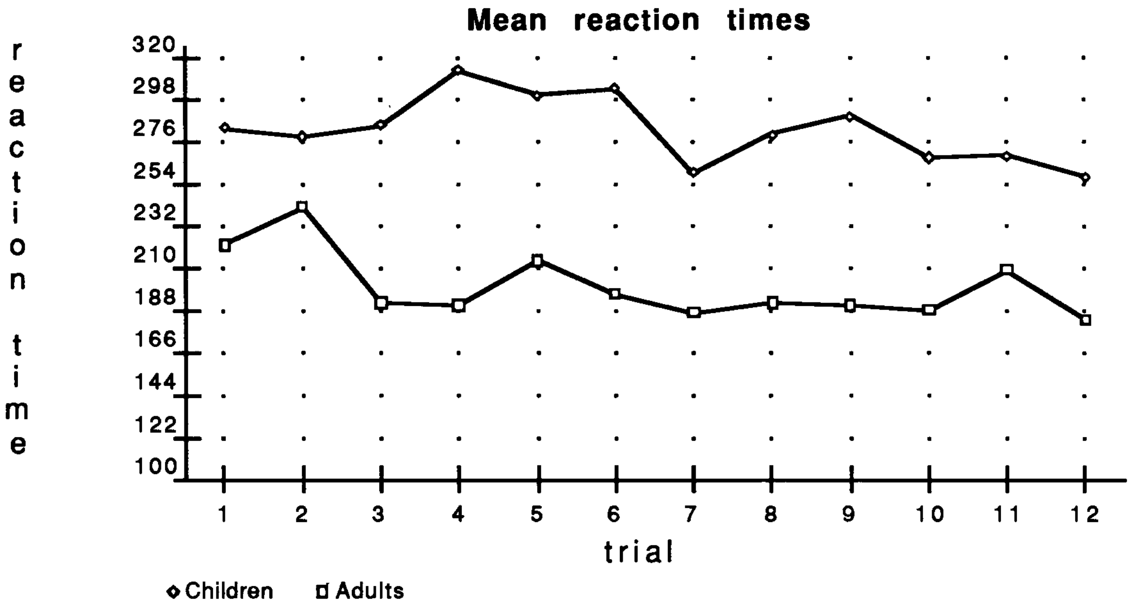
Mean RTs and Standard Deviations for CORRECT Responses

Condition	Trial	Mean RT (centi-seconds)	Standard Deviation (centi-seconds)
1. Same	1	222.3	67.7
	2	241.9	131.3
	3	191.8	35.8
	4	191.1	20.5
	5	214.1	46.3
	6	196.0	24.9
2. Different	7	187.2	20.2
(Same	8	191.9	27.1
Tonality)	9	190.4	20.2
3. Different	10	187.6	20.4
(Different	11	208.8	37.3
Tonality)	12	183.1	18.6

The mean RTs of the children and the adults are shown in Figure 4.2.

Figure 4.2

Mean RTs (centi-seconds) of both the children and adults for each trial



Variance–ratio tests of the mean response times for each subject's correct RT performance for each condition showed differences in variance between the different conditions. The variances of condition 1 (same stimuli) and condition 2 (different stimuli suggesting the same tonality) were significantly different (F(9,9)=6.68, *p*<0.005). The variability of response times in condition 1 (same stimuli) and condition 3 (different stimuli suggesting different tonality) were also significantly different (*F*(9,9)=5.47, *p*<0.01). However, the variability of RTs in condition 2 (different stimuli suggesting the same tonality) and condition 3 (different stimuli suggesting different tonality) were not significantly different ($F(9,9)=1.22$, not significant).

The mean RTs for the trials in each adult condition were also analysed by a Kruskal-Wallis ANOVA. However, unlike the children's responses, the Kruskal-Wallis ANOVA confirmed that there were no significant differences between the conditions ($\chi^2=4.33$, $p=0.115$, not significant).

The means and standard deviations for the three experimental conditions are shown in Table 4.5.

Table 4.5

Adult mean RTs and SD for each experimental condition

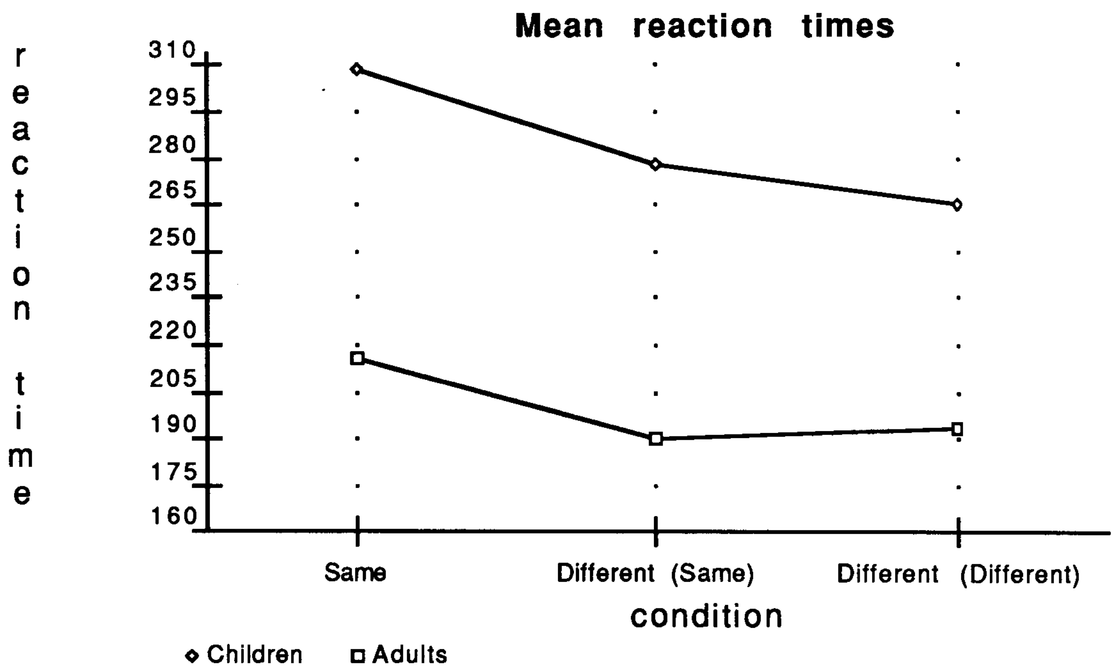
Condition	Mean RT	Standard Deviation
1. Same	215.8	50.7
2. Different (Same Tonality)	189.8	19.6
3. Different (Different Tonality)	193.2	21.7

The mean RTs for CORRECT responses to the twelve trials for both the children and adults were significantly different ($t(df=11)=12.22$, $p<0.0001$). A low, positive but non–significant correlation between the children's and adult's mean RTs for the twelve trials was observed ($r=0.138$, not significant).

The mean RTs of both the children and the adults for the three conditions are shown in Figure 4.3.

Figure 4.3

Mean RTs (centi-seconds) for both the children and adults for each condition



4.7. Discussion

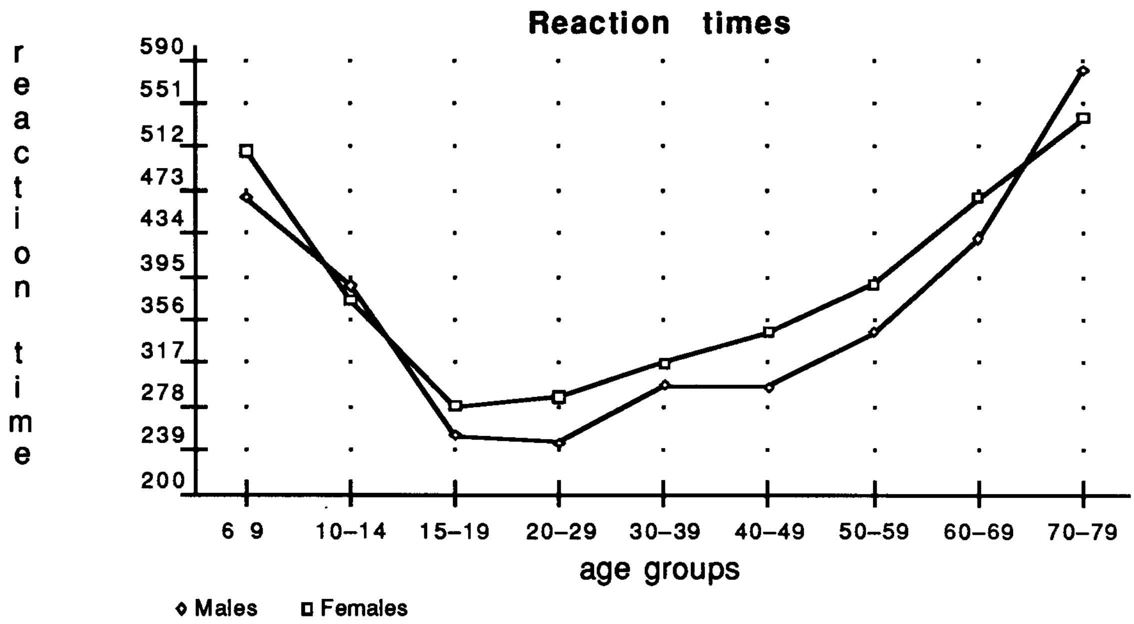
The lower error rate of the adults (2.8%) compared with the error rate of the children (17%) is consistent with the greater training of the musically experienced adults.

Consistent with previous research in other domains besides music, the RTs for the adult's responses to the experimental stimuli are much shorter than those of the children. For instance, both Hugin *et al.* (1960) and Hodgkins (1962) have found that simple RTs all shorten from childhood through adolescence to the twenties, followed by a slow increase in RTs until the sixties and a rapid lengthening in the seventies and beyond. Noble *et al.* (1964), who has undertaken one of the most detailed systematic investigations of the factors of both sex and age in choice RT, also recovered this profile. The RTs that Noble obtained from each age group are plotted in Figure 4.4.

Figure 4.4.

Mean RTs (milliseconds) for both males and females:

after Noble *et al.* (1964) from Welford (1980), *p.* 330



The shorter RTs for adults than those produced by children in the pilot study are consistent with the differences found by Noble *et al.* Noble also found that the males were generally faster than females, and that the differences between males and females were significant for all age groups except 10-14 and the oldest (aged over 70). Gender, however, is a factor not investigated by the pilot experiments.

As with the children, the mean RTs and standard deviations of the adults of the pilot study showed individual differences in variability of response times between the test conditions. However, adult subjects did not demonstrate a significant difference in discriminating between paired stimuli that suggested the same tonal schema and different tonal schema. Like the children, trial 12 of condition 3 (different stimuli suggesting different tonalities) possessed the shortest mean response time with the least variance. Although the shortest mean RT was for condition 2 (different stimuli suggesting the same tonal schema), this was not significantly different from condition 3 (different stimuli suggesting different tonal schemata). These results are compatible with the hypothesis that stimuli are initially abstracted to a tonal schema before processing of absolute or individual pitches ensues and that classification of same/different precedes such absolute pitch matching if different tonal schemata are suggested. The mean response times for condition 1 (same stimuli) were longer for both children and adults with considerable differences exhibited between subjects. It may well be that adults as well as children probably require mental rehearsal of the internalised representation of the same stimuli in order to discover any such changes.

The difference between the children and adults may be explained in that the tonal sense of musically experienced adults is so strong that tonality is disregarded by the adults in the comparison process. This would explain why the mean RT responses for each adult condition were not significantly different as revealed by a Kruskal–Wallis ANOVA, unlike the children's mean RT responses. An alternative explanation may be that the adults are able to separate the same/different character from the implied tonality of the stimuli to meet the criterion of the response instruction.

The longer RTs observed for correct responses of both children and adults to the *same* stimuli would seem to suggest an exhaustive search mechanism, although in no case could a response be made until after the final note of the stimulus had begun. This is consistent with the proposition that the processing of identical stimuli requires comparison along more dimensions and hence more time to effect a classification. Mental abstraction to a tonal schema would allow subjects to reject more quickly those stimuli which suggest different tonal schemata since the processing required to effect a response need not involve processing individual components of the stimuli but holistic matching of global attributes.

The significant differences in RT observed for the different conditions provide further support for the notion that different processing strategies are employed for pitch matching in different contexts. The fact that it takes comparatively less time to respond correctly to two *different* stimuli suggests that processing involves comparisons along a number of dimensions whereby a global abstraction to an internalised set of tonal relations precedes processing of absolute pitches. This accords with the model suggested by experiment one. The pilot experiments, therefore, lend support to the experimental hypotheses and provide further evidence for hierarchical processing. Those decisions which are made more quickly are hypothesised as operating as a fundamental process within the hierarchy. The diagrammatic representation of the model of hierarchical processing suggested by the previous pilot study is further supported by this experiment.

4.8. Implications of pilot experiments

The pilot experiments led to conclusions concerning the appropriateness of the methodology. The design of the next experiment was influenced by this pilot experiment which had explored *same* or *different* responses to stimulus pairs using intervals presented within the context of a tritone.

4.8.1. Sound quality

The pilot experiment used the internal sound chip of the BBC computer to present the sounds to subjects. The **S** and **D** keys on the computer keyboard were used by subjects to respond. For the next experiment, the sound quality of the stimulus signal was improved by connecting the BBC computer through a midi interface to an electronic keyboard from which the experimental stimuli were sounded.

4.8.2. Button-box for responses

Although the keys on the computer had been used in these pilot experiments, a custom built button box with two buttons would provide an easier response environment than two small keys on the computer keyboard. The computer was connected to an external button box with two buttons (one marked **S** for *same* and another marked **D** for *different*) which was used by subjects to indicate their response. This setup was subjected to an additional pilot experiment with a few children of varying ages to determine that the experimental method was satisfactory, although no data were analysed since too few responses were obtained.

Furthermore, this additional pilot experiment had revealed that subjects tended to use both hands to use the button box to respond to the experimental stimuli (usually the left hand for *same* and right hand for *different*). When some of the children in this additional pilot experiment were instructed to utilise one particular hand for both of the responses it slowed down the response time and caused confusion on the part of the respondent. A preference for the use of both hands for responses was considered appropriate in subsequent experiments, although it was recognised that this has implications for cerebral hemisphere effects.

4.8.3. Possible cerebral hemisphere–dominance effects

Auditory perception has been found to be affected by right or left brain hemisphere dominance (*e.g.* Kimura, 1961). Brain hemispheres are linked contralaterally to auditory pathways (*e.g.* left hemisphere-right ear) and the nervous system (*e.g.* left hemisphere-right hand). A shorter mean RT for a particular hand might indicate nothing more than particular hemisphere localisation and superiority for response to the task. A number of researchers have found evidence for cerebral hemisphere superiority for certain musical tasks. Kelley and Brandt (1984) found that the right hemisphere was generally more efficient in recognising pitch change than the left, with musicians producing a greater differentiation between the ears than non-musicians. However, Pechstedt, Kershner, and Kinsbourne (1989) found that musical training seemed to improve tonality processing in the left hemisphere. Their subjects demonstrated a right–ear advantage for discriminating between true transpositions of a melody and distractors than conserved tonality. Dichotic presentation of stimuli would be necessary to isolate an observed difference, although hemisphere specialisation is not an immediate concern in the experiments proposed here. Binaural presentation of stimuli seemed appropriate for this experiment in order to lessen any possible hemisphere superiority effect.

The mental translation of an aural stimulus into a linguistic response may be different for *same* and *different* responses. If this is suggested by differences in RTs in a simple discrimination task, then the significant results of the pilot experiments will need to be reconsidered. Moreover, hand superiority might be partially responsible for observed differences in RTs, particularly if processing is lateralised to a particular hemisphere. It was important to determine that the significant differences observed in the pilot experiments were not attributable to some secondary factor of the experimental method.

4.8.4. Practice effects

Another difficulty with the pilot materials had concerned the effect of practice. It was observed that RTs for the first few presented trials were longer than those of later trials, with decreasing times being observed after the first few trials of the experiment. While this experimental effect was negated to a certain extent by the randomised presentation of experimental trials in the pilot experiment, it was a very noticeable. Children seemed to be experiencing a short period of adjustment to the sounds of the stimuli themselves. This seemed to be overcome in the additional pilot experiment by a change to the experimental procedure which allowed each subject to be present in the room and listen to the previous subject's trials. However, subjects were not allowed to communicate with each other: the next subject sat on a chair unable to observe responses to the trials. This procedure, which had been adopted in the additional pilot experiment, was also adopted in the next experiment since it seemed to remove this experimental effect.

4.8.5. Experimental fatigue

The length of the block of experimental trials also seemed important to the children. In the first pilot experiment, no visual feedback was provided concerning the length of the experiment or the relative position within the trial block. In fact, a blank screen was initially presented to avoid visual distraction from the aural stimulus. However, subjects seemed to suffer from test fatigue in that they needed to know how many trials were going to be presented. Subjects in the additional pilot seemed much happier with the information concerning the total number of trials and the presentation of the trial number on the computer screen just before the trial was presented aurally. This incorporation of computer presented information concerning position within the experiment allowed the subjects to pace themselves within the block of trials in the next experiment.

4.8.6. Computer program amendments

Other amendments to the computer program were deemed desirable in the light of the pilot experiments. A teacher commented that the black and white nature of the instructions was unlike other programs which the children used regularly. The teacher suggested that the visual presentation of further experiments could be improved by the use of colour.

Considerable revision of the computer program used in the pilot experiment was necessary to ensure that unexpected responses by the children did not affect the data collection. The most important change ensured that inadvertent holding of the space–bar on the computer did not affect the results. The *auto-repeat* facility on most computers ensures that any key which is held down for more than half a second or so begins to send sequential keypresses at a very fast rate to a storage area (called a *buffer*) for later use. If subjects held down the space bar for any length of time during the inter–trial pause (which happened particularly with younger children), the multiple signals in the keyboard buffer generated through the auto-repeat facility allowed the program to shoot over inter–trial pauses and into the next trial. To overcome this problem, the buffer was cleared by the program every time a response was requested from the computer keyboard.

Quantification of the processing time for responses was measured by routines which reported the processing time taken to recognise that a button on the button–box had been depressed. The buttons were wired through the computer's analogue joystick fire button connectors, and the analogue to digital convertor was turned off by a software command. This improved processing time since the digital to analogue converter operates four independent software timers which were not needed. The operational delay was tested in the development stage by a software routine which reported in centi-seconds how much time the program needed to recognise the depression of the button. After considerable experimentation, the eventual delay was determined to be one centi–second, which is as small as can be expected. Some processing delay is inevitable since the program was written in the interpreted BASIC language which runs more slowly than a compiled language.

4.8.7. Summary

The changes to the experimental apparatus included improving the sound quality of the experimental materials and modifying the feedback mechanism with a button response box. The changes to the procedure involved familiarising the subjects with the experimental materials to negate the effects of practice and lengthening the trial block. The revisions to the computer programs were designed to improve their capacity to account for unexpected responses. All of these modifications were implemented in the next experiment reported in chapter five.

5. EXPERIMENT ONE: RESPONSES TO *SAME* OR *DIFFERENT* PAIRED–COMPARISON STIMULI

5.1. Rationale

5.1.1. Introduction

The first experiment investigated the discrimination of the interval of a semitone through forced–choice responses to *same* and *different* paired comparison stimuli. The collection of data to answer a number of important questions was sought by this particular *atomistic* procedure.

Like the pilot experiments, the experimental situation was conceptualised as a computer–driven closed environment in which the subject's interaction with the experimental materials would be managed by a computer. The computer system would present instructions, the stimuli of the experimental trials, measure the subject's RTs to the stimuli, and record the responses directly to a storage medium (floppy disk). Such a system ensured that the experimental conditions would be similar for all subjects. Moreover, possible errors in hand–recorded RTs were avoided by the computer's direct transcription to disk. Furthermore, subjects would interact with the experiment without human intervention: this would eliminate experimenter influence, which can be strong with children when tested individually. A computer–based system can give individual subjects control over the experimental environment, in particular the pace of presentation of the trials. It also permits interfacing to output devices (*e.g.* MIDI musical instruments) and input devices (*e.g.* switches, joystick and concept keypads), allowing a precisely controlled test environment.

Various currently available computer systems were investigated as possible environments which could have been used for the experiment. The advantages and disadvantages of the BBC microcomputer, Atari ST and Hybrid Music System were investigated. However, after review, the BBC computer was deemed to be the most suitable.[1]

A schema theory of music cognition acknowledges that cognitive abstraction is made possible by higher–order conceptual functioning. The assimilation of perceptual information to a schema affords a reduction in the information load of a stimulus. This characteristic of human cognitive behaviour is not peculiar to music. Although not strictly comparable to music cognition, recognition experiments using groups of letters which form nonsense words and meaningful words have demonstrated superior recall and faster response times for perceptually salient combinations of those groups which form higher–order cognitive units (*e.g.* Juola, Schadler, Chabot and McCaughey, 1978). The brain more readily makes sense of material if it can meaningfully be encoded to reduce the amount of information of the stimulus. In music cognition, the perceptual facilitation of the coding of redundancy within a recognised and practised cognitive structure (*e.g.* tonality) is developed

[1] A full discussion of the merits of each system is presented in Appendix 1.

largely by experience. It is this development with which the following experiments are principally concerned.

5.1.2. Hypotheses

The results of the pilot experiments using context–embedded semitones had demonstrated that shorter RTs were induced by stimuli which suggested different tonal centres. Accordingly, it was important to establish whether different mental processing strategies operate in the discrimination of *same* and *different* responses. This might be shown if a significant difference in RTs were observed for the two conditions of *same* and *different* stimuli in a context–free presentation.

The experiment described here was designed to evaluate the use of RT procedures with young children, and not to test their discrimination abilities. The experimental task was an uncontextualised simple pitch–matching task. If an uncontextualised forced–choice RT task generates significant differences between *same/different* conditions, then such differences could not be attributed to a context–generating prefix such as the tritone prefix used in the pilot experiments.

Some investigations of chronometric responses to musical stimuli have used a base–line response as a covariate in subsequent analysis. This covariate analysis has been used to control the wide subject differences found in RT studies. For example, Fiske (1982a) used a discrimination response to each of 25 beeps at various frequencies to calculate a mean

RT for each subject which he used as a covariate. The experiment reported here was designed to provide an indication of response times to paired comparison notes. Quantification of subject's RTs to a simple stimulus would give an indication of processing time needed for response and would subsequently be useful in comparing subject's responses to different types of stimuli, particularly if a covariate was required for later experiments.

This experiment was intended to illuminate the perceptual mechanisms of children in the primary age range (*i.e.* 7 to 11) related to the recognition of the interval of a semitone presented without the context–generating prefix used in the pilot materials. The experimental hypothesis was therefore that the dependent variables (*i.e.* error rates and RT for responses) would be affected by the independent variables (*i.e.* same/different stimuli, ascending/descending stimuli, and effects of both gender and age).

The null hypotheses of the experiment were:

i) no differences will be observed in correct classification for either of the two conditions (i.e. ***same*** *or* ***different****) as the responses in each category are equiprobable*

ii) no difference will be observed in RT responses for the various conditions of the experiment if the independent variables of type of stimuli and subject differences of age or gender are not exerting any effect.

The data obtained from the experiment comprised the *same* or *different* classification and the corresponding RT measurement. The number of correct or incorrect *same* or *different* responses determined error rates and produced nominal categorical data. The RTs produced data of an interval ratio nature, a higher level of measurement than categorical data.

5.1.3. Statistical Decisions

The binomial model (*cf.* Guildford and Fruchter, 1978, *p.* 186) is appropriate for the analysis of the proportions of nominal categorical data in *same* and *different* conditions. The relationship between error rates for *same* and *different* conditions and the independent variables would be investigated by correlation coefficients. A significant correlation between performance of the task and increasing age (perhaps related to increasing level of musical experience) may demonstrate developmental aspects of pitch perception, although it is recognised that general cognitive development could be responsible for certain observed experimental effects.

The pilot results had proved problematic concerning statistical analysis of RTs. Response times for *different* conditions produced a much larger standard deviation than did those for *same*. These differences were found to be statistically significant using the variance–ratio test for equal variance. This criterion of equal variance between two sets of data is assumed by both the *t*–test and ANOVA procedures. If the assumption of equal variance is violated by the experimental data there may be a necessity to use non–parametric procedures.

The experimental design necessitated some repetition of trials in order to include equal numbers of *same* and *different* trials. Since each *same* paired comparison was to be repeated it would be appropriate to ascertain the reliability of the test. A reliability coefficient (*i.e.* split half coefficient of reliability) could be calculated from the responses to the repeated stimuli.

5.1.4. Experimental Design

Subjects interacted with a BBC microcomputer environment in a repeated–measures design experiment. Both of the experimental conditions were tested within a single trial block. The computer determined a random order of presentation for each subject. This randomisation of presentation ensured that undesirable experimental effects resulting from greater experimental experience, such as increases in processing speed or improved accuracy of discrimination, would be evenly distributed between the trials.

5.2. Method

5.2.1. Subjects

Thirty–four subjects aged between six and eleven participated in the experiment.[2] This represented a span of five age–groups distributed over the age range as follows:

seven	subjects aged approx.	6– 7:	(4 boys, 3 girls)
ten	subjects aged approx.	7– 8:	(6 boys, 4 girls)
five	subjects aged approx.	8– 9:	(2 boys, 3 girls)
seven	subjects aged approx.	9–10:	(2 boys, 5 girls)
five	subjects aged approx.	10–11:	(4 boys, 1 girl)

There were 18 boys and 16 girls of varying musical backgrounds.

5.2.2. Materials and Apparatus

The apparatus for the experiment consisted of a BBC microcomputer system (model B) including keyboard, colour monitor and disk drive. The computer was linked to a K1 MIDI interface which was attached to a Yamaha PSS480 keyboard set to instrument number fifty–two (*i.e.* Piano 2) at fairly full volume. The button box was placed in front of the computer keyboard for easy access.[3]

[2] The primary age pupils attended a small two–class rural school in Derbyshire which was chosen since it allowed the experiment to be administered to a wide age range of children with the minimum of disruption. The pupils were known to the experimenter in teaching music to this class regularly over a term. .

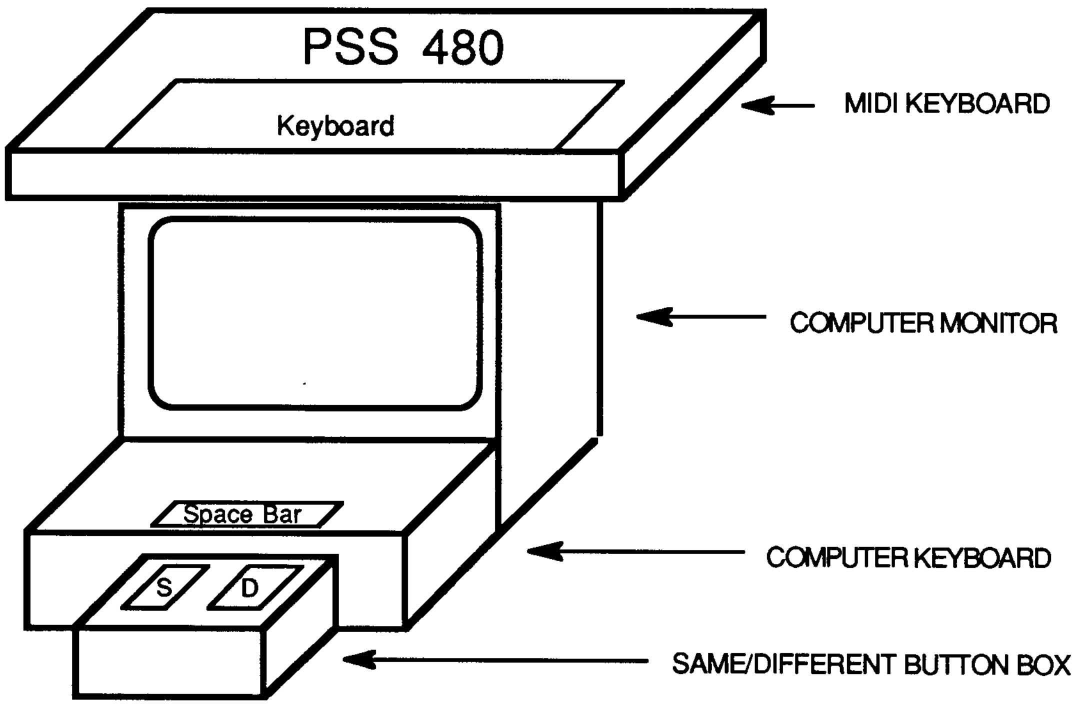
[3] Details of equipment are given in Appendix I.

Test stimuli comprised twenty trials: paired notes were identical in ten of these, and the remaining ten differed by a semitone.[4] All trials were given within a narrow pitch range deemed to be within the vocal range of the subjects (*i.e.* from middle C to F sharp). This range had been suggested as appropriate by Welch (1979,1989).

A schematic diagram of the experimental situation is given in Figure 5.1.

Figure 5.1

Schematic diagram of the experimental situation



[4] A musical representation of the materials is given in Appendix IV and the computer program which presented the trials in Appendix V.

5.2.3. Procedure

The subjects were tested individually in a room adjoining the classroom. The next subject to be tested was present in the room but was unable to observe the screen or interact with the subject. Subjects were all tested within the space of a week within the normal school day. Preliminary verbal instructions given to waiting subjects explained that they were present in the room to listen to the test, but they were not allowed to speak or look at the computer screen during the test. Following the observation of the previous subject, the subject sat in front of the computer and answered two questions verbally concerning the nature of the test. The first question asked the subject:

How many questions in the test?

Since subjects had heard the previous subject's answer to this question and had also listened to the experimental trials they found no difficulty in responding with the correct answer (*i.e.* twenty).

The second question asked the subject:

What do you have to do?

Subjects found no difficulty in explaining that they had to press **S** for *Same* or **D** for *Different* on the button box.

The experiment began with a title–page from the computer which scrolled the words:

Music Test

quickly across and down the computer screen in different colours simultaneously with a rapid glissando containing repeated notes sounded from the electronic keyboard. This had the dual effect of capturing the subject's attention and providing an opportunity for the experimenter to check that the volume setting of the keyboard was appropriate, and that the midi interface was functioning correctly. The instructions within the program began by requesting the subject's name:

WHAT IS YOUR NAME?

The subject responded by typing in his or her name, followed by pressing the RETURN key on the computer keyboard. The computer responded:

Pleased to meet you,

followed by the subject's name. The instructions were then presented in mode 7 double-height text in a variety of colours, but with important words contrasted in a different colour (indicated by underlining in the following).

YOU WILL HEAR TWO NOTES

was followed by the phrase

PRESS SPACE BAR TO CONTINUE

The instructions were separated by this command to press the space bar as the same process was used in the test to separate trials and it seemed sensible for this training to take place during the instructions since it encouraged the children to be independent of the experimenter during the test. The initial pilot experiment had caused some confusion with subjects occasionally turning to the experimenter between trials to ask what to do next and this process was avoided by the instructions being presented in this way. Whenever the space bar was pressed, the screen was cleared. The instructions continued with:

THEY MAY BE THE SAME

followed by

OR THEY MAY BE DIFFERENT

with the space bar pressed after each instruction.

PRESS 'S' IF THEY ARE THE SAME

PRESS 'D' IF THEY ARE DIFFERENT

PRESS SPACE BAR TO BEGIN

concluded the instructions.

The twenty experimental trials followed the instructions. The order of presentation for the twenty trials was revised by computer for each subject. After the opening screen, the inputting of the subject's name and instructions, each trial was preceded by a notification of the trial number appearing on the computer screen *e.g.*

Test Item Number 1

in a different random colour each time to give the visual appearance some variety. This was followed one second later by the first note of one second duration. After a silence of two seconds the comparison note was sounded for one second. As soon as the button was pressed the computer responded by confirming the button pressed with the message:

You Pressed SAME

or:

You Pressed DIFFERENT

depending on the response, even if the note had not finished sounding. This presented useful feedback to the subjects by showing that the computer had registered the response (and also allowed the experimenter to observe which response had been recorded).

The buttons on the box were colour coded (*i.e.* yellow for *same* and blue for *different*) and the capitalised words **SAME** and **DIFFERENT** were displayed on the computer screen in the appropriate corresponding colour. The RT of the paired standard/comparison stimulus was measured in centi–seconds from the beginning of the comparison note. This was the earliest that subjects could recognise the stimuli as different, and many subjects responded before the comparison note had sounded for one second. One second after the pressing of the appropriate response button, the subject was prompted at the bottom of the screen to:

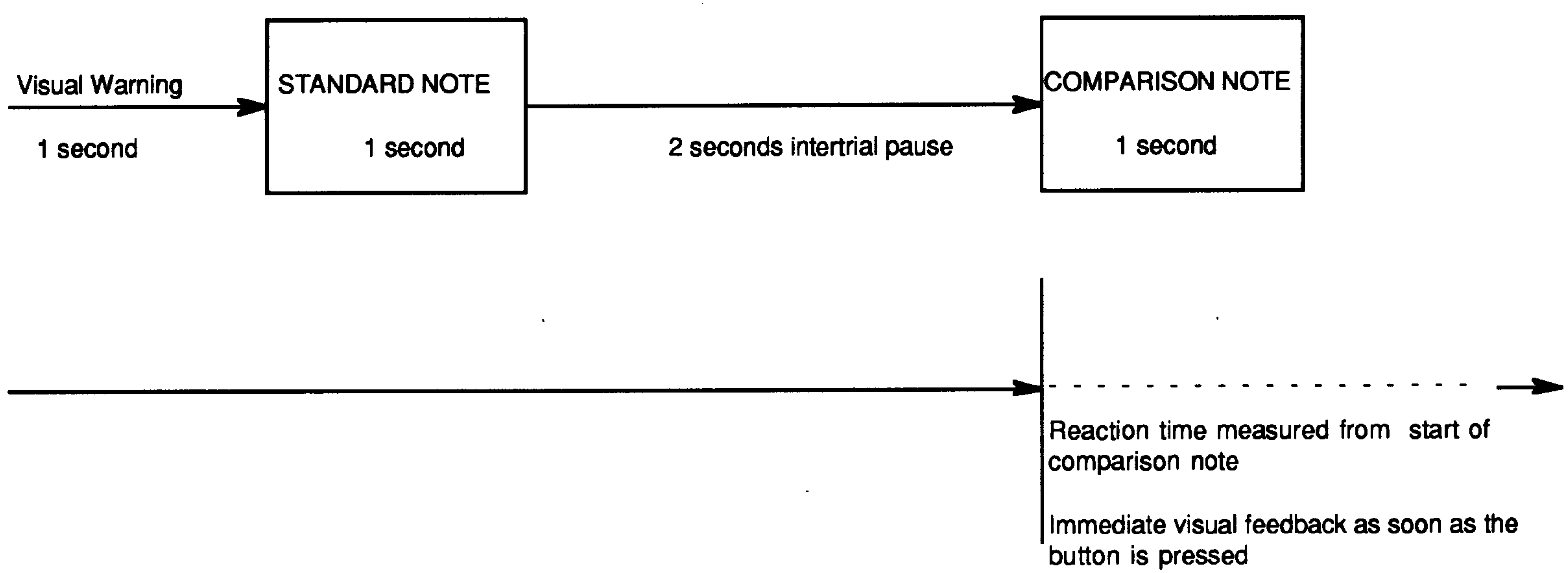
PRESS SPACE BAR TO CONTINUE

Thus the inter–trial pause was governed by the subject pressing the space bar to continue when ready.

A diagrammatic representation of the inter–stimulus time intervals displays the test procedure for each trial in Figure 5.2.

Figure 5.2

A diagrammatic representation of the inter–stimulus time intervals for each trial



After twenty items the program responded with

THANK YOU FOR YOUR HELP

while storing the response information on disk. The experiment concluded with a closing procedure which repeated the opening title screen involving rapid changing colours accompanied by a glissando on the midi instrument.

5.3. Results

5.3.1. Levels of Performance in the test

All thirty–four subjects seemed to enjoy the experiment. However, one subject seemed to find the cognitive demands of the test difficult. He answered every question as *same.* His results were discarded in the subsequent analysis.

5.3.1.1. Error responses for the total sample (aged 6-11)

Analysis of the responses from the total of 660 trials administered to the thirty–three remaining subjects showed that a total of 592 were answered correctly (90%) and that 68 were incorrect (10%). The distribution of responses to the *same* and *different* conditions is tabulated in Table 5.1.

TABLE 5.1

Incidence of CORRECT and INCORRECT responses to the *same* and *different* conditions

	SAME	DIFFERENT	TOTAL
CORRECT	323	269	592
INCORRECT	7	61	68
TOTAL	330	330	660

A chi–square test applied to these proportions produced a value of $\chi^2=46.05$ with 1 d.f. where $\chi^2=10.83$ was required for $p<0.001$ level. This is evidence for an association between *same/different* and CORRECT/INCORRECT responses.

5.3.1.2. Error responses for each age group

The number of CORRECT and INCORRECT responses for each year group is presented in Table 5.2. The number of errors is also expressed as a percentage to allow comparison as there were different numbers of subjects in each age group.

TABLE 5.2

Incidence of CORRECT and INCORRECT responses for each age group to the *same* and *different* conditions

Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial *p*
10–11	(5)	95	5	5	$p<0.0001$
9–10	(7)	135	5	4	$p<0.0001$
8–9	(5)	93	7	7	$p<0.0001$
7–8	(9)	160	20	11	$p<0.0001$
6–7	(7)	109	31	22	$p<0.0001$
	(n = 33)				

The distribution of *same/different* responses for each year group was evaluated by a binomial test and all were found to be highly significant (*i.e.* $p<0.0001$). All year groups were therefore differentiating those paired notes which were *different* from those which were *same*.

5.3.1.3. Error responses for 'same' condition

Since a significant difference in distribution of CORRECT/INCORRECT responses had been observed between the *same* and *different* responses, a further analysis of the incidence of error responses within each condition was undertaken.

The proportion of error responses made by each year group for the *same* condition is presented in Table 5.3 with associated probability values calculated by a binomial test. All year groups were discriminating correctly at a level significantly above chance (*i.e.* $p<0.0001$).

TABLE 5.3

Incidence of CORRECT and INCORRECT responses in the *same* condition

Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial *p*
10–11	(5)	49	1	2	$p<0.0001$
9–10	(7)	70	0	0	$p<0.0001$
8–9	(5)	49	1	2	$p<0.0001$
7–8	(9)	88	2	2	$p<0.0001$
6–7	(7)	67	3	4	$p<0.0001$
	TOTAL	323	7		
	(n = 33)				

5.3.1.4. Error responses for 'different' condition

The proportion of error responses made by each year group for the *different* condition is presented in Table 5.4 with associated probability values calculated by a binomial test. All year groups were responding appropriately except the youngest age group, *i.e.* six and seven year old pupils.

TABLE 5.4

Incidence of CORRECT and INCORRECT responses in the *different* condition

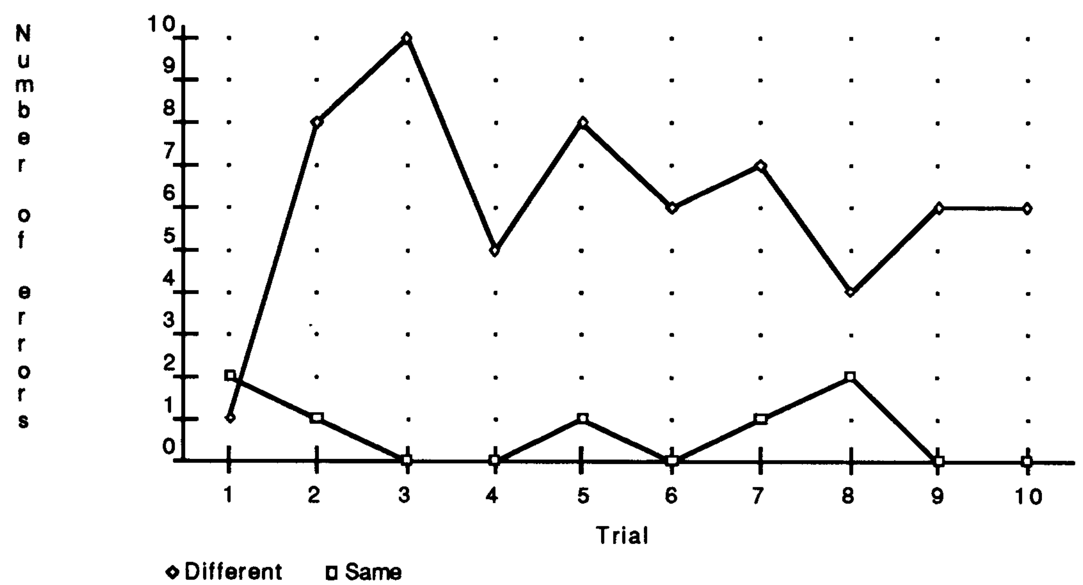
Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial *p*
10–11	(5)	46	4	8	$p<0.0001$
9–10	(7)	65	5	7	$p<0.0001$
8–9	(5)	44	6	12	$p<0.0001$
7–8	(9)	72	18	20	$p<0.0001$
6–7	(7)	42	28	40	non–sig.
	TOTAL	269	61		
	(n = 33)				

5.3.1.5. Error responses for matched trials

Having established that the number of INCORRECT *different* condition responses (*i.e.* 61) was significantly different from the INCORRECT *same* condition responses (*i.e.* 7), the distribution of the error responses according to trial was examined. The distribution of error responses for the matched trials for each condition of *same* and *different* is represented in Figure 5.2. Matched trials are those with identical standard stimuli.

Figure 5.3

Number of error responses for each matched *same* and *different* trial



Although some variability of error responses across trials was clearly evident, a chi–square goodness–of–fit test to a rectangular distribution assessed whether error responses were equally distributed between the trials of the ten different condition error responses.[5] This produced a value of $\chi^2=9$ with 9 degrees of freedom which failed to reach the $\chi^2=16.92$ critical value required for $p<0.05$ level. There was therefore no

[5] The rectangular distribution assumes the equal probability of occurrence of different alternatives.

evidence to suggest significant departure from equal distribution of error responses for experimental trials within the *different* condition. There were so few *same* condition error responses that a similar analysis for *same* condition responses was unnecessary. Subjects were therefore finding equal difficulty with all trials within the *different* condition in responding accurately to the stimuli.

5.3.1.6. Age group differences

The number of error responses made by each year group is presented in Table 5.5.

TABLE 5.5

Incidence of *same* and *different* error responses

Age Group	(n)	Same Errors	Different Errors	Total Errors	Mean Error
10–11	(5)	1	4	5	1.00
9–10	(7)	0	5	5	0.71
8–9	(5)	1	6	7	1.40
7–8	(9)	2	18	20	2.22
6–7	(7)	3	28	31	4.43
Total	(33)	7	61	68	2.06

This distribution shows that increasing age was inversely correlated with decreasing error judgments (*rho*= –0.9, N=5; $p<0.05$). This is a clear indication that younger pupils found either the discrimination exercise or the management of the response system difficult. Whether their difficulty was due to cognitive problems or the demands of coordinating motor output is uncertain.

5.3.1.7. Gender differences

The subject sample contained unequal numbers of boys and girls in the 9–11 age group. The 10–11 year group (n=5) possessed four boys and one girl (who made no errors) whereas the 9–10 year group (n=7) possessed two boys and five girls. The error rates for both boys and girls for each year are given in Table 5.6.

TABLE 5.6

Error rates for both boys and girls distributed by year

Year Group	Boys' Errors	Boys (n)	Girls' Errors	Girls (n)	Mean Boys' Errors	Mean Girls' Errors
10–11	4	(4)	0	(1)	1.00	0.00
9–10	1	(2)	4	(5)	0.50	0.80
8–9	4	(2)	3	(3)	2.00	1.00
7–8	13	(5)	7	(4)	2.60	1.75
6–7	17	(4)	14	(3)	4.25	4.67
Totals	39	17	28	16	2.29	1.75

These mean error responses for each year group are significantly positively correlated (*rho*= –0.9, N=5; p<0.05) between the boys and girls performances. Although the mean error rate obtained by boys is higher than that obtained by girls, the two sets of mean scores did not demonstrate significantly different means computed by a Mann–Whitney U–test, indicating that differences in gender are less important than differences in age.

5.3.2. Reaction time responses

Data were analysed by subject mean RT, trial mean RT, and age group mean RT. Means were generated from either ALL responses or from CORRECT responses only.

5.3.2.1 Subject analysis of RTs for ALL trials

Mean RTs for each subject, expressed in centi–seconds, for each of the two conditions are presented in Table 5.6.[6]

[6] The computer programs for this computation are presented in Appendix V

TABLE 5.7

Mean RTs for ALL responses of each subject under the two conditions (*i.e. same* and *different*)

Subject	Age group	Mean RT of ten *same* trials (centi-seconds)	Mean RT of ten *different* trials (centi-seconds)
1	10-11	128.6	155.4
2	10-11	129.2	106.2
3	10-11	126.0	114.8
4	10-11	163.6	179.5
5	10-11	182.8	182.2
6	9-10	98.3	108.6
7	9-10	148.7	155.5
8	9-10	174.7	206.4
9	9-10	186.4	169.6
10	9-10	131.3	134.0
11	9-10	161.9	210.2
12	9-10	160.6	158.0
13	8-9	126.0	131.9
14	8-9	138.5	174.2
15	8-9	127.8	135.3
16	8-9	165.4	160.5
17	8-9	140.3	174.2
18	7-8	105.4	112.5
19	7-8	159.1	185.7
20	7-8	161.2	163.5
21	7-8	143.6	183.0
22	7-8	202.1	184.7
23	7-8	151.2	194.3
24	7-8	166.6	248.5
25	7-8	132.7	124.9
26	7-8	144.0	104.3
27	6-7	170.4	226.2
28	6-7	247.7	229.1
29	6-7	199.3	195.5
30	6-7	145.0	235.7
31	6-7	261.8	703.1
32	6-7	223.1	239.3
33	6-7	261.7	207.2
Mean =		162.6	187.7
Standard Deviation =		40.6	101.2

Table 5.7 shows that the overall mean RT for the *same* stimuli was shorter than that for *different* stimuli. A variance–ratio test of the two conditions of *same* and *different* was highly significant ($F(32,32)=6.21$; $p<0.01$). However, a correlated *t*–test applied to means proved not significant.[7]

The significant variance–ratio test confirmed a finding of the earlier piloted materials in that the responses to *different* items demonstrated greater variability in RT than those for *same* items. Some pupils reacted much more quickly to certain *different* stimuli than to other *different* stimuli. This may be attributable to different processing strategies used by different pupils or to factors within the stimuli. The non–significant difference in means between the two conditions suggests that the distinction between *same* and *different* paired comparisons was not distinct perceptually or cognitively. However, the variability of the *different* condition trials seems to suggest differences in processing strategies between the conditions of *same* and *different* . Two strategies are suggested by this data. Subjects may hold an image of the first stimulus in memory until the match is confirmed by the comparison stimulus. Alternatively, subjects could hold the image of the first stimulus in memory until contradicted by the second stimulus, checking back on the stimulus image to make comparison. This is consistent with Piaget's notion of reversibility.

[7] Although the homogeneity of variance requirement of parametric tests such as the *t*–test was violated by the data, since the data being compared were taken from the same sample with correlated trials, a correlated *t*–test was appropriate (*cf.* Guildford and Fruchter, *p.* 159).

The RTs between the two conditions were highly positively correlated (r=0.649; associated t value = 4.76 ; p=0.0001). This shows that subjects were responding in a systematic manner to both of the conditions, and confirms that similar RTs were being elicited for both the *same* and *different* conditions.

5.3.2.2 Subject analysis of RT responses for CORRECT trials

A similar analysis of RTs of the CORRECT responses was also undertaken. Error responses, by their very nature, can generate longer RTs since they can result from indecision following cognitive confusion and may be guesses. As a statistically significant difference in error responses had already been demonstrated between the two conditions, the inclusion of error response RTs might bias the results. The means of the reactions times for correct decisions only are presented in Table 5.8.

TABLE 5.8

Means RTs for CORRECT responses of each subject under the two conditions (*i.e. same* and *different*)

Subject	Age group	Mean RT of *same* (centi–seconds)	Mean RT of *different* (centi–seconds)
1	10-11	128.6	147.0
2	10-11	128.6	106.2
3	10-11	126.0	113.4
4	10-11	163.6	182.4
5	10-11	182.8	182.2
6	9-10	98.3	108.6
7	9-10	148.7	146.9
8	9-10	174.7	206.4
9	9-10	186.4	169.6
10	9-10	131.3	134.0
11	9-10	161.9	205.9
12	9-10	160.6	148.6
13	8-9	125.2	120.4
14	8-9	138.5	131.1
15	8-9	127.8	138.8
16	8-9	165.4	158.2
17	8-9	140.3	173.3
18	7-8	105.4	111.6
19	7-8	159.1	184.9
20	7-8	161.2	147.1
21	7-8	143.6	184.0
22	7-8	182.9	185.9
23	7-8	151.2	193.0
24	7-8	166.6	217.0
25	7-8	132.7	124.9
26	7-8	144.0	107.0
27	6-7	170.4	226.2
28	6-7	247.7	170.5
29	6-7	182.8	231.0
30	6-7	145.0	320.7
31	6-7	261.8	602.3
32	6-7	223.1	197.0
33	6-7	260.2	195.1
Mean		161.4	180.9
Standard Deviation		39.6	88.2

The mean correct RT of 161.4 centi–seconds in Table 5.8 is similar to the mean of 162.9 centi–seconds reported in Table 5.7. Moreover, the standard deviations of the *same* condition in the two tables are very similar (*i.e.* 40.6 in Table 5.7 and 39.6 in Table 5.8). However, the *different* condition, with its many more error responses removed, showed a decrease in RT from 187.7 centi–seconds for ALL responses to 180.9. Similarly, the standard deviation also decreased from 101.2 to 88.2. This reduced difference in means and decreased variability was further examined and significant differences in variance were still observed ($F(32,32) = 4.96$; $p < 0.01$).

The decreased variance and smaller mean of the different condition obtained by removal of INCORRECT responses would suggest that much of the variance is attributable to the error responses. This is compatible with the notion that the delay is caused by cognitive processing or mental rehearsal of the stimuli preceding a wrongly–guessed error response. This might be attributable to *internal noise* (Krueger, 1978; cf. chapter two). However, the significant differences in variance between the *same* and *different* responses is not accounted for by the many more error responses for the *different* condition alone. Moreover, the fact that a consistent pattern of response to the conditions is lacking might indicate that different cognitive processing strategies are being used by different subjects. While certain subjects responded more quickly to the *different* condition than to the *same* condition, others seemed to adopt a different profile of responding. Some cognitive factor seems to be present in the discrimination response times, but appears to be attributable to subject differences, rather than to context of the experimental materials.

However, the RTs for subjects across the two conditions were highly positively correlated (r = 0.601; associated t value = 4.2 ; p=0.0003), indicating a high level of subject consistency in response behaviour.

5.3.2.3 Trial analysis of RTs for ALL trials

The RTs were also examined for outliers which also might have biased the results and some of the 6–7 year old subjects were observed to have produced greater variability in RTs than the older children. Examination of the data showed that much of the variability within the ten trials of the *same* condition was being lost by comparing only the means of each condition for each subject. The calculation of the means for each subject's performance under each of the two conditions reduced the variability which could be seen in RTs for each trial. The analysis of trials might reveal perceptually salient features of the experimental materials. The data were re–analysed by ascertaining means and standard deviations for each trial, since the simplistic global classification into *same* and *different* might not be the perceptually salient characteristic of the stimuli to which subjects were attending.

The mean RTs for each trial and standard deviations of RTs are given in Table 5.9.

TABLE 5.9

Mean RTs and SD
for ALL experimental trials

Trial Number		Mean RT	SD
1	(same)	153.5	39.7
2	(same)	156.3	71.7
3	(same)	173.0	86.8
4	(same)	180.8	95.0
5	(same)	161.4	49.4
6	(same)	152.4	37.8
7	(same)	168.4	66.1
8	(same)	167.0	55.7
9	(same)	153.9	41.8
10	(same)	159.1	62.6
11	(different)	180.0	89.1
12	(different)	197.6	283.2
13	(different)	176.8	99.8
14	(different)	183.9	65.3
15	(different)	179.2	91.7
16	(different)	168.8	61.3
17	(different)	185.0	99.4
18	(different)	242.1	339.7
19	(different)	178.2	84.1
20	(different)	185.5	116.1

Mean (of ALL *same* RTs) = 162.58 centi–seconds
Standard Deviation (of ALL *same* RTs) = 9.48 centi–seconds

Mean (of ALL *different* RTs) = 187.71 centi–seconds
Standard Deviation (of ALL *different* RTs) = 20.50 centi–seconds

Trial analysis preserved the mean RT recovered by subject analysis for each condition.[8] However, the reduction to ten scores for each condition resulted in a reduction of the amount of variability demonstrated by the standard deviations. The Standard Deviation of 40.53 for all the *same* condition subject responses has reduced to a Standard Deviation of 9.48 for responses of the ten *same* conditions. Similarly, the Standard Deviation of 101.16 for the *different* subject condition has become 20.5 for the ten *different* analysis of trials. This shows that the effect of subject variability is reduced by an analysis from mean scores which are summated across trials. However, two of the *different* trials (trials 12 and 18) produced large deviations.

Again a variance ratio test revealed that the two conditions were significantly different ($F(9,9) = 4.68$; $p<0.05$). The RTs for all of the experimental responses were subjected to a correlated *t*–test between the ten *same* and ten *different* responses and the result was found to be highly significant ($t(9)= -3.52$; $p<0.01$). A non–parametric Wilcoxon matched–pairs signed ranks test between the *same* and *different* mean RTs confirmed this significant difference (T=0 with N=10; $p<0.01$). The means for the trials of the two conditions are therefore significantly different. This is to be expected, given the significant differences in error rates between the two conditions. It might be that slower response

[8] Slight differences are the result of rounding errors caused in part by the number of significant figures carried by the computer for the calculations.

times of the error responses which occur more frequently in the different condition are responsible for much of the significant difference in means.

Although the means were different, the two sets of mean RTs exhibited a very low positive correlation (r= +0.196; associated t value=0.56, p=0.59, not significant). There seems, therefore, no linear relationship or interaction between the RTs of the two conditions.

The interaction of RT and errors was investigated by computing the correlation between the RTs and the number of errors for each trial (r = +0.44 ; associated t value = 2.08; p<0.05). This significant result confirmed that that the higher RT responses were significantly linked with increased error rates. The *same* responses were faster and induced fewer errors than *different* responses.

5.3.2.4. Trial analysis of RTs for CORRECT trials

The data were also re–analysed with only those RTs which were CORRECT. The results are shown in Table 5.10.

TABLE 5.10

Mean RTs and SD for CORRECT responses to each experimental trial

Trial Number		Mean RT	SD
1	(same)	148.0	30.5
2	(same)	150.1	63.2
3	(same)	173.0	86.8
4	(same)	180.8	95.0
5	(same)	157.8	45.7
6	(same)	152.4	37.8
7	(same)	169.4	66.9
8	(same)	165.1	54.6
9	(same)	153.9	41.8
10	(same)	159.1	62.6
11	(different)	180.1	90.5
12	(different)	141.0	30.5
13	(different)	155.9	64.0
14	(different)	174.6	60.9
15	(different)	172.2	103.5
16	(different)	155.7	56.1
17	(different)	174.3	102.2
18	(different)	181.6	362.8
19	(different)	163.9	44.7
20	(different)	156.3	53.4

Mean (of CORRECT *same* RTs) = 160.96 centi–seconds
SD (of CORRECT *same* responses) = 10.80 centi–seconds

Mean (of CORRECT *different* RTs) = 165.56 centi–seconds
SD (of CORRECT *different* responses) = 13.14 centi–seconds

The RT data were examined for outliers, since large RTs (which would indicate indecision or inattention) might unduly bias the results. One outlying RT of over 20 seconds for trial eighteen, caused by a disturbance to the experimental situation, was removed from the analysis. A variance ratio test yielded a non–significant result ($F(9,9) = 1.48$; not significant) satisfying the homogeneity of variance requirement for parametric analysis.

A correlated *t*–test between the *same* and *different* CORRECT response times gave a non–significant result between the two observed means of 160.96 and 165.56 ($t(9) = -1.01$; $p=0.34$, not significant). The correlation observed between the RTs for the two conditions was low and positive, but was not significant ($r=+0.293$; associated t value =0.87, $p=0.58$). The significant difference in means observed in the analysis of all responses can therefore be attributed to the error responses, which generally manifested longer response times. A two–way ANOVA on data with error scores removed found no significant differences between either the conditions or the trials.

The analysis of correct RTs shows that the means for the two conditions were not significantly different. Furthermore, they were not correlated. This lack of correlation indicated that no linear relationship existed between the RTs of the two conditions, and confirmed that the matched stimuli were not inducing systematic differences in RTs.

5.3.2.5 Age group analysis of RTs for CORRECT trials

The relationship between age and RT was investigated by an ANOVA of the means of CORRECT responses for each trial for each year group. The calculated values are tabulated in Table 5.11.

TABLE 5.11

Mean RTs of CORRECT responses
for each trial for each age group

Ages	10–11	9–10	8–9	7–8	6–7
	(n=5)	(n=7)	(n=5)	(n=9)	(n=7)
Trial					
1	147	133	162.8	144	159.5
2	140.8	133.4	138	130.4	217
3	147.6	157.1	143.6	148.1	260.1
4	141.4	172	140.2	145.4	292
5	161.8	144	134.4	154.2	195.7
6	161.6	139.1	129.6	158.3	167.6
7	164.75	157.7	142.2	148.3	230.3
8	131.2	175.9	155	166.9	182.6
9	135.6	163	125.2	148.1	186
10	134.6	141.7	129.4	146.3	231.6
11	129.8	137.3	178.5	175.9	265.1
12	127.25	148.2	135.3	137.3	156
13	144.6	174.4	107.8	146.3	208.3
14	161.25	144.8	149	165.3	242.8
15	147.5	218.7	150	135.9	211
16	140.6	150.2	136.2	149.3	241.7
17	204	135.4	139	168.3	298
18	125.2	158.8	143.8	201.8	266.2
19	160.8	157.3	167.5	151.9	196.8
20	138.2	161.3	145.8	142	225.7
Mean	= 147.28	155.17	142.67	153.2	221.7
SD	= 18.18	20.09	15.62	16.22	41.46

A one–way analysis of variance of the mean RTs of correct responses for each trial for each year group was computed and the differences across age groups was found to be significant (Table 5.12).

TABLE 5.12

ANOVA of mean RTs of CORRECT responses
for each trial for each age group

ANALYSIS OF VARIANCE TABLE (ONE–WAY)

SOURCE	S.S	DF	MS	MSR	*p*
Ages	85176.2	4	21294.10	35.9687	*p*>0.00005
Residual	56241.6	95	592.02		
TOTAL	141417.8	99			

The computed value of F is very highly significant ($F(4,95) = 35.9687$; $p<0.00005$). However, the greatest variance (shown by the standard deviation of 41.46 for the 6–7 age group) tested against the smallest variance (15.62 for the 8–9 age group) was significant ($F(19,19) = 7.04$; $p<0.01$). A Tukey HSD test confirmed that the 6–7 year old group was significantly different from all other age groups at a $p<0.05$ level.[9] This analysis of year group RTs invited the postulation that this variability might be a function of less developed cognitive processing strategies of younger pupils.

[9] The Tukey HSD (Honestly Significant Difference) test is particularly conservative.

The correlation between the various year groups of mean RTs of responses for all trials is shown by the correlation matrix (Table 5.13). No correlations were significant, indicating the absence of a linear relationship between the RT responses of the various year groups.

TABLE 5.13

Correlation matrix of mean RTs of CORRECT responses for ALL trials for each age group

Ages	10–11	9–10	8–9	7–8	6–7
10–11	1.000	–.213	–.052	0.022	0.230
9–10		1.000	–.078	–.167	–.016
8–9			1.000	0.268	0.094
7–8				1.000	0.388
6–7					1.000

5.4. Discussion

The primary purpose of this experiment was to evaluate RT procedures with young children, particularly exploring the discrimination task of semitone discrimination. The experiment confirms the findings of the pilot experiments.

5.4.1. Discrimination

Junior school children (aged 7-11) were able to discriminate the interval of a semitone in this matched paired–note context. Moreover, children as young as six (*i.e.* top infants) were able to discriminate semitones at a significant level. The utilisation of the semitone as a perceptually salient discrimination interval across the age range six to eleven is thus justified in the experiments reported here. The uneven distribution of errors between the *same* and *different* conditions suggests that discrimination is less salient for the condition which compares *different* notes. It may be that the mechanism underlying the internalisation and coding of pitches which are the same is less dependent on cognitive processing than comparison of different pitches.

5.4.2. Age and error responses

Older children certainly performed better than younger children in making fewer error responses. The significant inverse correlation of increasing age and decreasing error judgments suggests a developmental effect. It is hypothesised that this effect is a product of the cognitive demands of the test, and it may be that older children can internalise the given pitch at a deeper level in the perceptual hierarchy as they possess more developed tonal schemata. The test difficulty might be a factor which is influencing the number of errors, but as the procedure for responding to the *same* and *different* conditions was identical, some alternative explanation seems necessary to explain the observed differences in error rates between the two conditions of *same* and *different*.

5.4.3. Effects of musical experience

The effects of musical experience were considered. Although some younger children had recently become involved in extra–curricular instrumental lessons, the teaching programme was not sufficiently advanced to enable investigation of the effects of musical training on the experiment to be quantified explicitly. Some children who performed accurately were receiving extra music lessons on an instrument, but it was noted that some of the younger children who made more errors were also receiving extra instrumental lessons. Owing to the small number of subjects, further analysis was not possible.

5.4.4. RTs of *same* and *different* responses

The non–significant difference of the mean RTs between the two conditions of the responses of the subjects does not suggest that the cognitive processes required to differentiate between *same* and *different* notes in isolation are distinct. However, the observed significant difference in variance indicates that some children found the processing of *different* notes more difficult than others, suggested by the larger standard deviations for the *different* condition. An important uncontrollable effect might be significant here, particularly the contextualisation of a response in relation to the previously heard trials. It was intended that the randomisation of the presentation of the trials would negate any progressive practice effect.

The significant correlation between the conditions of *same* and *different* is further evidence that there was a consistent relationship in the time taken by subjects to respond to stimuli across the two conditions.

The mean RTs of the two conditions were found to be significantly different when derived from ALL responses to the experimental trials. However, the means calculated with the RTs for error responses removed (*i.e.* CORRECT responses only) were found to be non–significant. This confirms that the error responses were responsible for the difference in means. This suggests that subjects were, in fact, recognising the difference intuitively, but were unable to confirm this difference by a cognitive process and so responded in error after a processing delay. This supports the hypothesis that RT is indicative of an hierarchical cognitive processing model. The processing seems to be distinct for *same* and *different* responses, although within each condition the non–significant differences would seem to suggest similar processing models.

The analysis of the mean trial RT CORRECT responses confirmed that the error responses were responsible for much of the observed differences in means and differences in variability. The lack of linear correlation between the *same* and *different* matched trials also confirms an absence of systematic variability. In fact, this lack of correlation confirms no perceptual similarity between the matched trials, *i.e.* the pitches were not matched perceptually and cognitive abstraction was unrelated to the instantiation of a tonal schema.

5.4.5. Age and RT responses

The RTs produced by the youngest age group (i.e. 6–7 year old) were revealed as significantly different from other age groups by the ANOVA of mean RTs of CORRECT responses of trials for each year group. This might be attributable to a different processing mechanism being utilised by the younger children. However, this difference might be partially attributable to the classroom musical experiences of the infant children (aged 6–7) being different from the musical experiences of the junior children (aged 7–11).[10] This difference in musical experiences may be an important factor which is partly responsible for this observed significant difference in variability in RTs, although it was not possible to investigate

[10] The class teacher of the junior children was an accomplished musician who regularly involved her pupils in music lessons, whereas the infant teacher admitted her reluctance and lack of experience of teaching music. This suspicion was confirmed by the musical performances of the two classes: the infant children did seem to sing particularly poorly, certainly compared with the older children.

this systematically The lack of a significant difference between the other age groups as revealed by the Tukey HSD test provides no evidence to suggest that different processing strategies were being utilised by children in the 7–11 age group. Alternatively, it might be that younger children were taking longer to respond and being more variable in their response times as they were using different processing models, or that simply the translation of an aural stimulus to a verbal discriminatory response took longer in younger subjects. A possible simplification of the cognitive demands of the experiment could involve a yes/no response to the question *'are the stimuli the same?'*. However, all subjects clearly understood the concepts of *same* and *different*, and the variability and longer response times were for the *different* condition. The nature of the stimulus seems a much more pertinent variable which is affecting the response time, although it is recognised that observed effects might have alternative explanations related to other cognitive factors apart from aural processing models.

5.4.6. Conclusion

One of the primary purposes of the experiment was to investigate handedness and whether any differences in processing time could be attributed to the experimental method. The lack of significant difference with the uncontextualised stimuli of this experiment confirms that the method is appropriate for measuring choice RT and is not unduly affected by experimental method. Any significant difference would have necessitated a reversal of the right hand/left hand buttons on the button–box for half the sample. The results of this experiment obviate the need for this cautionary measure.

The lack of any significant correlation between the mean correct RT responses of age groups for the two conditions does suggest that the comparison of two pitch neighbours does not demand higher–order abstraction of structural features. Longer stimuli than those used in this experiment are necessary to instantiate abstraction to a tonal schema. Higher–order cognitive functioning would more likely result from the greater information load of longer stimuli and consequent interference effects between pitches.

The cognitive processes of interest to the research question are those of the deeper level of the cognitive hierarchy, presuming that a hierarchical processing model is the most fitting explanation for the observed differences in RT. A further experiment, specifically devised to test the effect of the contextualisation of the semitone in relation to other pitches could reveal greater differences of processing time than the isolated non–contextualised presentation of two pitches and therefore give an indication of the level of nesting in the hierarchy. Any observed differences in a contextualised presentation would have greater significance as the non–contextual presentation has confirmed no relationship in RT responses between the conditions of *same* or *different*, although one developmental factor has been identified.

EXPERIMENT TWO: RESPONSES TO *SAME* OR *DIFFERENT* PAIRED–COMPARISON NOTES IN MAJOR TRIADIC CONTEXT

6.1. Rationale

6.1.1. Introduction

Hypotheses concerning the discrimination of the interval of the semitone with subjects of primary schools age were tested in experiment one. In experiment two the same stimuli as experiment one were contextualised within a specific tonality by utilising a triadic prefix in conjunction with a suffix note. The experiment was therefore an extension of the previous experiment. The same procedure and method were repeated with the same subjects, using materials modified to involve a paired comparison of a sequence of four notes rather than two single notes. Although this increased the length of the stimulus and correspondingly the duration of the experiment, more information could be extracted concerning the effects of context on the prefix and suffix.

This experiment aimed to establish whether a prefix such as the tonic triad is a sufficient stimulus for defining a tonal context for children. This is claimed by a number of researchers such as Cuddy and Badertscher (1987) and Trehub (1987). Cuddy and Badertscher (1987) found that the diminished triad, with its tonally specific tritone interval component, was an insufficient context–generator to recover the tonal hierarchy as obtained by Krumhansl. This question is of crucial importance in determining whether children abstract and assimilate pitches to a tonal schema using the intervallic rivalry of less frequent intervals proposed by Butler (1989).

The previous experience of the subjects could influence this second experiment as shorter observed RTs might result from the effects of practice. Alternatively, longer stimuli could take more processing time and RTs might correspondingly increase.

The longer stimuli of the second experiment could affect error rates. These longer stimuli would possibly need to be abstracted at a deeper level in the proposed hierarchy in order to facilitate a comparison. Younger children with a less developed perceptual coding mechanism could find abstraction more difficult and consequently make more errors. Observed differences in error rates or RTs could indicate a different processing strategy being employed from that used in the first experiment. Whereas short–term memory might suffice for the first experiment in a direct comparison of two notes with a relatively short inter–note time interval, processing to long–term memory might be necessary to process the longer four–note stimuli. However, the triadic prefix was repeated for each stimulus and consequently would soon become familiar to subjects. Again, this experiment required a randomised presentation order for all trials to distribute any processing hesitation inevitable with test familiarity.

6.1.2. Hypotheses

The primary interest in this experiment was accordingly the effect of the contextualisation of trials and the corresponding influence that this would have on either the number of errors or the speed of RT responses.

This experiment investigated the perceptual mechanisms of primary school–age children in discriminating the interval of semitone in the contextual presentation of triadic prefixes. The experimental research question explored how the same dependent variables as the previous experiment *(i.e.* error rates and RTs) would be affected by the independent variables now presented within a triadic contextual prefix (*i.e. same* and *different* stimuli, gender and age).

The null hypotheses were that:

i) there will be no significant difference in the distribution of the equiprobable classificatory responses of same and different .

ii) there will be no significant difference in mean error rates observed in different age ranges or between boys and girls.

iii) the mean RTs observed between the conditions of same and different will not be significantly different if subjects are using similar processing *strategies for each condition.*

iv) the mean RTs for the correct only responses will not be significantly different between the conditions of same or different

v) the mean RTs for the correct only responses will not be significantly different between the trials

vi) the RTs should exhibit no significant degree of correlation between the matched trials of the two conditions

Significant differences would indicate that the contextualisation is affecting the responses.

6.1.3. Experimental Design

The repetition of *same* condition trials within the experimental design (required to balance the number of trials within each condition of *same* and *different*) could indicate the consistency of the test. This analysis was not undertaken in the previous experiment as a *prima facie* examination of RTs did not seem to suggest a high level of correlation, although with the additional information provided by this experiment, a number of analyses were considered appropriate. For instance, a test of the internal consistency of RTs in this second experiment would be indicated by the correlation of the *same* trials which are duplicated. Furthermore, a comparison between the consistency of this experiment and the previous experiment could measure the correlation between the mean RTs of the twenty trials of each experiment.

The grouping of stimuli was also recognised as an important area of analysis. The perceptually salient features of trials could be revealed by examining the musical characteristics of those trials which seem to possess an association between RTs or error responses. Any attempt to reveal cognitive hierarchical processes by means of quantifying differences in RTs required detailed analysis of the characteristic features of the trials themselves.

Certain problems of interpretation still remained in this experiment. The two conditions of *same* and *different* in the previous experiment exhibited significant differences in variance. This polarisation of variability of RTs between the two conditions could be exacerbated by the longer stimuli. However, this experimental effect, if observed, would suggest different processing strategies at work, and could be attributable, as in the previous experiment, to the delays preceding incorrect responses.

The design of experiment two was identical to that of experiment one. Subjects interacted individually with a BBC microcomputer in a repeated measures design with subjects experiencing both of the experimental conditions of *same* and *different* within the same trial block. The computer determined a randomised order of presentation for each subject.

6.2. Method

6.2.1. Subjects

The subject sample was taken from the same classes as the previous experiment. The thirty–three subjects representing five age–groups comprised 18 boys and 15 girls. The subjects were the same as the previous experiment with two changes: year six (10–11 year olds) included one extra subject and year two (6–7 year olds) excluded one subject. The subject who was unable to understand the cognitive demands of experiment one and whose results had previously been discarded did not undertake this experiment.

6.2.2. Materials

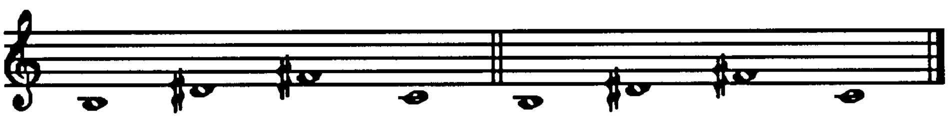
The trial materials were an extension of the previous experimental paired–notes which were either the *same* or *different* by a semitone. A triadic prefix was placed before each of the notes of the original experiment.

The ten *same* and ten *different* condition stimuli were considered to be matched if the standard stimuli were identical. Within experiment two, four of the *same* condition stimuli were repeated as *different* condition stimuli involving both ascending and descending semitone changes to the comparison suffix note. Since the range of the paired notes was C to F, the context–defining prefix chosen was the triad of B major (*i.e.* B, D sharp and F sharp), whose range encompassed all suffix notes. The triad was prefixed before each note of the comparison.[1] An example of one trial of each of the experimental conditions is given below (Figure 6.1).

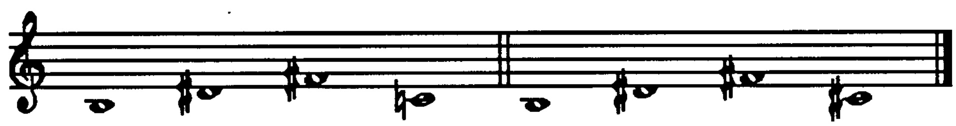
Figure 6.1

Musical example of *same* and *different* stimuli

Musical example of *same* stimulus



Musical example of *different* stimulus



[1] A musical representation of the absolute pitches of the experimental trials is presented in Appendix VI and the computer program to present these pitches is given in Appendix VII.

6.2.3. Procedure

The subjects were tested individually in the room adjoining the main classroom with the next subject present in the room. The junior class subjects (aged 7–11) were tested within a period of one week within the normal school day and the infants (aged 6–7) some three weeks later. This followed within a six–week interval from the first experiment. The preparatory verbal questioning of the subjects was the same as the previous experiment and the instructions were identical except that they were instructed that:

YOU WILL HEAR TWO 4–NOTE TUNES.

The procedure was similar to the previous experiment but the paired comparison consisted of a four–note stimulus instead of the paired single notes of the previous experiment. The duration of each note of the triadic prefix was 750 milliseconds followed by the first suffix note of one second. After a silence of two seconds, the three triadic prefix notes of 750 milliseconds each were followed by the comparison suffix note of one second.

The method of measurement of RTs and procedure were the same as the previous experiment.

6.3. Results

6.3.1. Levels of performance in the test

Performance levels were analysed in the same way as experiment one, investigating errors by condition, year groups, matched trials, age group and gender.

6.3.1.1. Error Responses for the total sample (aged 6–11)

The number of CORRECT responses fell from 592 (90%) of the last experiment to 485 (73%) in experiment two. The distribution of responses by condition is tabulated in Table 6.1.

TABLE 6.1

Incidence of CORRECT and INCORRECT responses
to the *same* and *different* conditions

	SAME	DIFFERENT	TOTAL
CORRECT	272	213	485
INCORRECT	58	117	175
TOTAL	330	330	660

A chi–square test was applied to these proportions and yielded a value of χ^2 =26.16 with 1 *d.f.* where 10.83 was required for p<0.001 indicating that the proportions of responses were significantly different between the two experimental conditions (*i.e. same* and *different*). This is clear evidence for departure from chance responses.

6.3.1.2. Error Responses for each age group

The number of CORRECT and INCORRECT responses for each year group is presented in Table 6.2.

TABLE 6.2

Incidence of CORRECT and INCORRECT responses for each age group to the *same* and *different* conditions

Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial *p*
10–11	(6)	98	22	18	$p<0.0001$
9–10	(7)	113	27	19	$p<0.0001$
8–9	(5)	70	30	30	$p<0.0001$
7–8	(8)	119	41	26	$p<0.0001$
6–7	(7)	85	55	39	$p<0.01$
	(n = 33)				

The significance of the distribution of responses between observed and equiprobable expected values was evaluated by a binomial test. The unequal proportions of all year groups were found to be significant (*i.e.* $p<0.01$), although the proportion of younger children performing incorrectly was greater than older children. The 7–11 year olds were clearly discriminating between those stimuli which were the *same* and those which were *different*, but the 6–7 year olds found the discriminatory nature of the experiment, or the response mechanism, more difficult than older children. Subjects in the total sample were therefore clearly discriminating the interval of a semitone in this contextual presentation.

6.3.1.3. Error Responses for 'same' condition

Analysis of the incidence of error responses within the *same* condition by a binomial test revealed that all year groups were discriminating responses correctly at a highly significant level (Table 6.3).

TABLE 6.3

Incidence of CORRECT and INCORRECT responses in the *same* condition

Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial p
10–11	(6)	52	8	13	$p<0.0001$
9–10	(7)	57	13	19	$p<0.0001$
8–9	(5)	40	10	20	$p<0.0001$
7–8	(8)	66	14	18	$p<0.0001$
6–7	(7)	57	13	19	$p<0.0001$
	(n = 33)				

6.3.1.4. Error Responses for 'different' condition

The number of error responses made by each year group for the *different* condition is presented in Table 6.4. Year groups were discriminating significantly except for 8–9 year olds and 6–7 year olds.

TABLE 6.4

Incidence of CORRECT and INCORRECT responses in the *different* condition

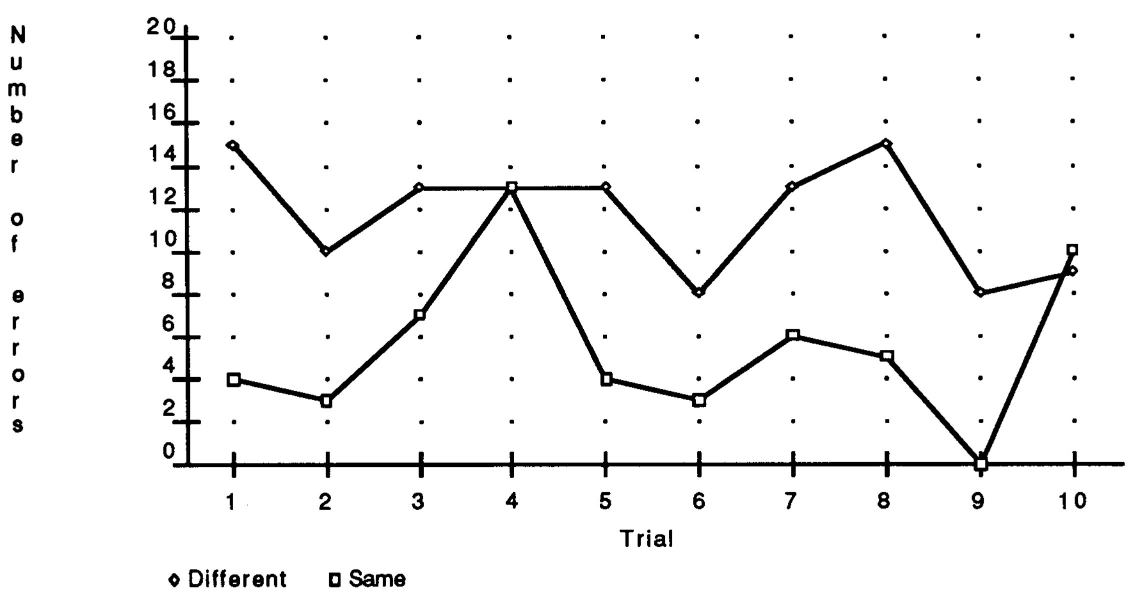
Age Group	Subject (n)	Number Correct	Number Incorrect	% Incorrect	Binomial *p*
10–11	(6)	46	14	23	$p<0.0001$
9–10	(7)	56	14	20	$p<0.0001$
8–9	(5)	30	20	40	not sig.
7–8	(8)	53	27	34	$p<0.001$
6–7	(7)	28	42	60	not sig.
	(n = 33)				

6.3.1.5. Error Responses for matched trials

The number of error responses for each trial was examined in a similar manner to the first experiment, *i.e.* by a chi–square goodness–of–fit test to a rectangular distribution examined whether error responses were equally distributed between the trials.[2] The number of error responses for each of the twenty trials is given in Figure 6.2. The *same* and *different* trials in which the suffix notes of the standard stimulus were identical are plotted together.

Figure 6.2

Number of error responses for each matched trial



[2] The *rectangular distribution* assumes the equal probability of occurrence of different alternatives.

A chi–square test of goodness–of–fit to the rectangular distribution of the twenty trial error responses produced a value of χ^2=39.06 with 19 d.f. which was larger than the χ^2=30.14 critical value required for $p<0.05$ level. There is therefore confirmatory evidence of significant departure from equal distribution of error responses for experimental trials within both the *same* and *different* conditions.

The *different* condition trial error responses were subjected to a chi–square goodness–of–fit test and the result was found to be non–significant (*i.e.* χ^2 = 5.65 (9 d.f.); p=0.78, not significant). This was congruent with the finding in the previous experiment and confirmed that the errors made for *different* trials were evenly distributed between the trials.

However, a similar chi–square goodness–of–fit test to a normal distribution analysis of the *same* condition error responses gave a significant value of chi–square (*i.e.* χ^2 = 17.52 (9 *d.f.*); $p<0.05$). Trials four and ten both had a larger number of errors than other trials, accounting in all for 40% of the total errors for same responses.

The suffix for the fourth trial (*i.e.* D natural) was the minor third of the tonic implied by the triadic prefix (*i.e.* B natural) and the suffix for the tenth trial was a tritone higher (*i.e.* F natural) than the tonic implied by the triadic prefix (*i.e.* B natural). This is shown by the musical example of Figure 6.3.

Figure 6.3

Experimental materials for trials four and ten

The confusion caused by the suffix for the fourth trial, a minor third above the implied tonic suggested by the major triadic prefix, is surprising. However, it might be that the incongruous minor third as part of the stimulus is confused in the respondent's mind with the *different* response. A number of older, more musically experienced children, (*e.g.* 50% of 10–11 year olds) failed to respond correctly to the stimulus of trial four.

An analysis of the characteristics of those trials which generated more error responses suggested that more errors were generated by suffix notes of the standard stimulus which are outside the diatonicism of the tonic suggested by the prefix. This does not, however, adequately explain the larger number of error responses for trial ten. The tonal ambiguity of the tritone in relation to the suggested tonicalisation of trial ten also seemed to create more errors. Trial ten encompasses a tonally-consistent wider range of notes than other trials in including the tritone and two other notes which unequivocally specify the tonality, although this tonality is F sharp (or G flat), the dominant of the tonic B suggested by other trials. Subjects may not have interpreted the tritone as establishing a different tonal centre in the experiment. An experimental effect of implied harmonic stability may have been created by the repetition of the triadic prefix. This repetition would establish the root of the triad over successive trials and consequently define the tonality.

6.3.1.6. Age group differences

The number of error responses made by each year group is presented in Table 6.5. The mean subject error, calculated by dividing the total number of errors for each age group by the number of subjects, gives an indication of the higher number of errors for younger subjects for the twenty trials of the experiment.

TABLE 6.5

Incidence of *same* and *different* error responses

Age Group	Subject (n)	*Same* Errors	*Different* Errors	Total Errors	Mean Subject Error
10–11	(6)	8	14	22	3.67
9–10	(7)	13	14	27	3.86
8–9	(5)	10	20	30	6.00
7–8	(8)	14	27	41	5.13
6–7	(7)	13	42	55	7.86
Total	(33)	58	117	175	5.30

The pattern of increasing age being inversely correlated with decreasing error judgments observed in the first experiment is sustained by these mean error rates. The correlation between all year groups and mean error rates gave a negative value of *rho*= –9 with N=5; $p<0.05$. This significant correlation is similar to that observed in the first experiment. However, the mean subject errors observed in this experiment are significantly greater than those of the first experiment ($t(4)= -9.96$, $p=0.001$).

The significant correlations confirm that younger pupils found the discrimination more difficult. The significant difference in means suggests that all pupils found the longer contextualised stimuli more difficult to discriminate than the uncontextualised stimuli.

6.3.1.7. Gender differences

The judgments of older children were examined as there were unequal numbers of boys and girls in each year. The 10–11 age group (N=6) possessed five boys and one girl whereas the 9–10 age group (N=7) possessed two boys and five girls. The error rates for both boys and girls for each year are given in Table 6.6.

TABLE 6.6

Error rates for both boys and girls distributed by year

Year Group	Boys Errors	Boys (n)	Girls Errors	Girls (n)	Mean Boys Errors	Mean Girls Errors
10–11	19	(5)	3	(1)	3.80	3.00
9–10	5	(2)	22	(5)	2.50	4.40
8–9	19	(2)	11	(3)	9.50	3.67
7–8	30	(5)	11	(3)	6.00	3.67
6–7	32	(4)	23	(3)	8.00	7.67
Total	105	(18)	70	(15)	5.83	4.67

Unlike the previous experiment, where the mean error responses for the boys and girls of each year were significantly positively correlated, the correlation between the performances of the boys and girls observed in this experiment does not approach significance (*rho*= +0.102 with N=5; *p*=0.435). Although the mean error rate obtained by boys is again higher than that obtained by girls, the two sets of mean scores did not demonstrate significantly different means computed by a Mann–Whitney U–test.

The relatively small sample sizes of each year group may be partially responsible for this observed effect of girl superiority, or alternatively, the girls might have more musical experience. As the children had similar classroom musical experiences, it would be surprising to find a significant difference between the abilities of the boys and girls to discriminate semitones in this contextual presentation. However, the lack of correlation is surprising. Perhaps the superior performance of the girls is telling us something about the cognitive abilities of the subjects related to their previous musical experiences and corresponding disposition to respond to the experimental materials. An examination of the reading ages on the Suffolk Reading Scale (1981), which the teacher had obtained with these children, showed that the older girls generally possessed higher reading ages than the older boys. This confirms that the two groups were not matched, suggesting that the lack of correlation is attributable to other factors apart from gender difference. These subject differences were explored more systematically in the next experiment.

6.3.2. Reaction Time Responses

Like experiment one, data were analysed by subject mean RT, trial mean RT and age group RT. Means were generated from either ALL responses or from CORRECT responses only.

6.3.2.1. Subject analysis of RTs for ALL trials

The mean RTs for all the responses of each subject in centi–seconds for each of the two conditions of *same* and *different* are presented in Table 6.7.

TABLE 6.7
Mean RTs for ALL responses of each subject
under the two conditions (*i.e. same* and *different*)

Subject	Age Group	Mean RT of ten *same* trials (centi–seconds)	Mean RT of ten *different* trials (centi–seconds)
1	10–11	127.7	127.0
2	10–11	105.7	108.7
3	10–11	149.6	152.2
4	10–11	133.9	124.0
5	10–11	159.2	178.6
6	10–11	164.8	194.0
7	9–10	94.8	99.4
8	9–10	144.6	177.8
9	9–10	340.9	185.0
10	9–10	151.8	191.7
11	9–10	145.8	132.4
12	9–10	173.7	181.2
13	9–10	167.8	158.5
14	8–9	173.5	171.6
15	8–9	110.8	122.3
16	8–9	138.0	151.4
17	8–9	215.7	138.8
18	8–9	206.3	185.7
19	7–8	122.4	116.6
20	7–8	167.0	143.5
21	7–8	160.9	132.7
22	7–8	226.5	200.4
23	7–8	162.6	216.9
24	7–8	184.2	240.8
25	7–8	114.0	128.3
26	7–8	137.2	163.6
27	6–7	174.4	246.7
28	6–7	153.7	168.2
29	6–7	130.7	194.2
30	6–7	165.7	183.4
31	6–7	394.0	303.1
32	6–7	140.8	157.4
33	6–7	266.8	215.5
	Mean =	169.86	169.44
	SD =	62.36	43.80

The contextual presentation of the stimuli produced remarkably consistent RT means for each of the two conditions (*i.e.* 169.86 for *same* and 169.44 for *different*). Interestingly, the significant difference observed in the non–contextualised presentation of the previous experiment (*i.e.* 162.6 for same and 187.7 for different) was not sustained in this experiment. The high positive correlation observed between the *same* condition and *different* condition RTs was preserved (r = +0.689; associated t value = 5.29; $p<0.0001$). This demonstrates that the contextualising influence was eliciting similar responses from subjects to the two conditions. The significant correlation indicated by this measure of internal consistency in the experiment suggests a degree of reliability concerning the RTs.

A two–way ANOVA without replications (*i.e.* where the residual variance includes interaction effects) reported a significant effect for subjects ($F(32) = 4.68$; $p<0.0001$). However, this result must be interpreted in the context of a variance–ratio test for related (or correlated) variances (Bruning and Kintz, 1977, *p.* 110) which gave a highly significant probability ($t(31) = 3.6$; $p<0.01$). The greater standard deviation for the *same* condition of experiment two (*i.e.* 62.36) than that observed in experiment one (*i.e.* 40.6) was somewhat surprising: however, analysis of the data revealed that subjects nine and thirty–one both exhibited comparatively long RTs for the *same* condition. The smaller standard deviation for the *different* condition of experiment two (*i.e.* 43.80) than that obtained in experiment one (*i.e.* 101.2) was also unexpected. This reduction in standard deviation would be compatible with the assumption that the compared notes were easier to discriminate in the contextualised presentation than the non–contextualised.

6.3.2.2. Subject analysis of RTs for CORRECT trials

The means of the RTs for the CORRECT responses are presented in Table 6.8.

TABLE 6.8

Mean RTs for CORRECT responses of each subject under the two conditions (*i.e. same* and *different*)

Subject	Age–group	Mean RT of *same* (centi–seconds)	Mean RT of *different* (centi–seconds)
1	10–11	131.8	127.0
2	10–11	107.2	96.3
3	10–11	152.6	153.8
4	10–11	133.9	120.6
5	10–11	159.2	164.4
6	10–11	155.6	179.5
7	9–10	94.8	101.0
8	9–10	133.3	177.8
9	9–10	220.3	158.9
10	9–10	154.1	143.6
11	9–10	134.5	132.4
12	9–10	183.0	186.7
13	9–10	160.3	164.1
14	8–9	172.4	164.7
15	8–9	94.7	88.7
16	8–9	135.2	151.4
17	8–9	199.3	123.4
18	8–9	206.4	169.5
19	7–8	131.4	134.1
20	7–8	172.1	138.0
21	7–8	167.3	122.7
22	7–8	223.3	207.4
23	7–8	162.6	185.2
24	7–8	157.8	236.6
25	7–8	114.0	144.2
26	7–8	137.1	157.4
27	6–7	174.4	255.0
29	6–7	141.5	165.3
30	6–7	165.7	159.3
31	6–7	266.1	220.2
32	6–7	128.7	128.8
33	6–7	251.3	226.8
	Mean =	160.1	158.9
	SD =	40.8	39.7

One of the subjects (subject 28 in year two) failed to achieve any correct answers to the *different* condition: in fact, he answered every trial with the *same* response. This meant that a mean value for the *different* condition could not be calculated. His results for both *same* and *different* conditions were therefore completely removed from the subject analysis as it was assumed that he had misunderstood the requirements of the experiment.

The means and standard deviations for the two conditions were found to be very similar. The contextualisation of the suffix notes in this second experiment produced no significant differences in means and standard deviations which were similar to those previously observed in the first experiment. Thus the removal of error response times demonstrated less effect than in the previous experiment.

A highly significant correlation between the two conditions was also observed ($r= +0.647$, associated t value = 4.65; $p< 0.0001$). This confirmed the finding that subject's RT responses were consistently reliable. Subject differences were further analysed by a two–way ANOVA (without replications). This showed significant differences for subjects ($F(31,31) = 4.67$; $p< 0.001$), but not for the conditions of *same* or *different* ($F(1,31) = 0.0375$, not significant). The ANOVA Table is presented in Table 6.9.

TABLE 6.9

Two–way ANOVA of mean RTs for CORRECT responses of each subject

ANALYSIS OF VARIANCE TABLE (TWO–WAY)

SOURCE	S.S	DF	MS	MSR	*p*
Reaction Times	21.5	1	21.51	0.0375	not sig.
Subjects	82870.1	31	2673.23	4.6616	p<0.001
Residual	17777.1	31	573.46		
Total	100668.7	63			

It is recognised that attempting to generalise about absolute durations of RTs from these data is difficult since subjects exhibited significant differences in response times. However, the increased length of the stimuli did not seem to induce a corresponding lengthening of response time. On the contrary, the mean RT and variability of response time as indicated by the standard deviation was smaller in this experiment for *different* responses than in the previous experiment. However, the much closer correspondence between means and standard deviations of the two conditions of the second experiment indicated that the contextualisation produced similar processing delays for each condition.

6.3.2.3. Trial analysis of RTs for ALL trials

The analysis of trials examined the mean RTs and standard deviations for ALL the experimental trials. The data are presented in Table 6.10.

TABLE 6.10

Mean RTs and SD for ALL experimental trials

Trial Number	Condition	Mean RT	SD
1	Same	167.2	100.7
2	Same	151.0	45.5
3	Same	218.7	230.3
4	Same	168.9	98.4
5	Same	176.8	93.2
6	Same	143.8	82.7
7	Same	160.2	143.6
8	Same	171.5	67.7
9	Same	178.9	86.2
10	Same	161.6	52.1
11	Different	185.1	111.1
12	Different	167.6	92.2
13	Different	190.5	95.7
14	Different	165.8	63.6
15	Different	165.2	55.3
16	Different	153.1	68.8
17	Different	163.7	74.5
18	Different	179.5	67.2
19	Different	162.3	64.9
20	Different	161.6	60.3

Mean (of ALL *Same* RTs) = 169.86 centi–seconds
SD (of ALL *Same* RTs) = 20.34 centi–seconds

Mean (of ALL *Different* RTs) = 169.44 centi–seconds
SD (of ALL *Different* RTs) = 11.72 centi–seconds

One interesting difference between RTs for the contextualised and non–contextualised stimuli of this experiment is the absence of the significant difference between the means found in the first experiment. A Wilcoxon matched–pairs signed–ranks test had demonstrated significance of $p<0.01$ in the analysis of trials of the previous experiment. The second experiment's contextualised stimuli were not significantly different (169.86 for the *same* condition and 169.44 for the *different* condition). This lack of difference was confirmed by a non–significant value of t ($t(9)= -0.169$; $p= 0.864$) in a t–test for correlated means.

A two–way analysis of variance confirmed that although the different overall means for each condition were non–significant ($F(1,9) = 0.01$; $p=0.928$, NS), the RTs for each trial were significantly different ($F(9,9) =4.03$; $p=0.025$). The ANOVA Table is presented in Table 6.11.

TABLE 6.11

Two–way ANOVA of mean RTs and SD for ALL trials

ANALYSIS OF VARIANCE TABLE (TWO–WAY)

SOURCE	S.S	DF	MS	MSR	*p*
Condition	0.88	1	0.88	0.01	not sig.
Trials	3973.29	9	441.48	4.03	0.025
Residual	985.96	9	109.55		
TOTAL	4960.13	19			

In fact, the two conditions seemed to exhibit a much closer correspondence in observed RTs than the first experiment. The low positive correlation of the first experiment was non–significant (r= +0.196; associated t value = 0.56; p= 0.59, non significant). However, the same paired comparison notes of the second experiment contextualised by a triadic prefix engendered RTs which were significantly positively correlated (r= +0.696; associated t value = 2.74; p= 0.025).

It must be concluded that this significant positive correlation of reaction times was attributable to the presence of contextual prefix notes, and it provided evidence of their effect on cognitive processing.

The larger standard deviation observed in this experiment for the *same* condition (*i.e.* SD = 20.34) than for the *different* condition (*i.e.* SD = 11.72) is the reverse of the data obtained in the first experiment. In that situation, the *same* condition demonstrated smaller variability of RTs. The respective variances of these two conditions could be compared by a variance–ratio test for correlated samples to give an indication of the variability of the RTs in each of the two conditions, but since a significant difference in error rates in each condition has already been established, the differences might be attributable to delays induced by error responses.

6.3.2.4. Trial analysis of RTs for CORRECT trials

An analysis was made of mean RTs and standard deviations for CORRECT only responses for each experimental trial in order to investigate the effect of the error responses. The means and standard deviations are presented in Table 6.12.

TABLE 6.12

Mean RTs and SD for CORRECT responses to each trial

Trial Number	Condition	Mean RT	SD
1	Same	162.4	93.4
2	Same	152.8	46.1
3	Same	185.7	121.7
4	Same	157.8	68.6
5	Same	171.1	92.4
6	Same	132.3	42.1
7	Same	130.3	37.7
8	Same	171.0	70.5
9	Same	162.9	64.8
10	Same	156.1	48.2
11	Different	168.7	61.2
12	Different	174.3	105.1
13	Different	181.0	101.2
14	Different	155.2	47.6
15	Different	176.7	60.4
16	Different	143.0	61.3
17	Different	140.5	36.2
18	Different	166.7	69.8
19	Different	139.4	42.8
20	Different	143.8	45.2

Mean (of CORRECT *Same* RTs) = 158.24 centi–seconds
SD (of CORRECT *Same* RTs) = 17.03 centi–seconds

Mean (of CORRECT *Different* RTs) = 158.93 centi–seconds
SD (of CORRECT *Different* RTs) = 16.37 centi–seconds

The means for the CORRECT responses only, although reduced from those obtained from ALL responses, were almost identical for each condition (*i.e.* 158.24 for the *same* condition and 158.93 for the *different* condition). Furthermore, the standard deviations are almost equivalent (*i.e.* 17.03 for *same* and 16.37 for *different*). The greater variability observed in the analysis of all the responses was removed with the removal of the INCORRECT responses. This contrasts with the results of the first experiment where a difference in means and standard deviations was observed.

The CORRECT response RTs between the conditions of *same* and *different* in the first experiment exhibited a low positive correlation (r=+0.293; associated t value = 0.87; p=0.58). However, the contextualisation of the matched trials of the second experiment produced a significant positive correlation in RTs (r= +0.702; associated t value= 2.79; p=0.023).

This significant correlation confirms that the contextual prefix was responsible for the observed relationship in reaction times and provides evidence for the perceptual reality of the mental abstraction of stimuli to a tonal schema.

The observed differences in trials were analysed by a two–way ANOVA. No significant difference was observed between the conditions of *same* and *different* (F(1,9) = 0.03; p=0.864, not significant) of the second experiment. However, differences in observed RTs between the trials themselves of the second experiment were found to be highly significant ($F(9,9) = 5.7$; $p=0.009$). These significant differences between trials were not found in the non–contextualised presentation of the previous experiment ($F= 1.36$; $p= 0.33$, not significant). The ANOVA Table for this second experiment is presented in Table 6.13.

TABLE 6.13

Two–way ANOVA for mean RTs for CORRECT responses to each trial

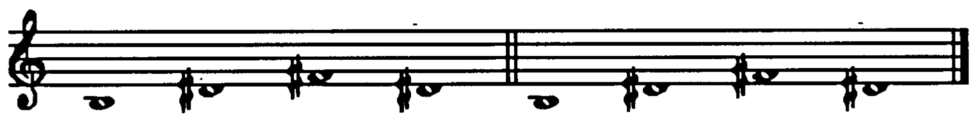
ANALYSIS OF VARIANCE TABLE (TWO–WAY)

SOURCE	S.S	DF	MS	MSR	*p*
Condition	2.38	1	2.38	0.03	not sig.
Trials	4272.79	9	474.75	5.70	0.009
residual	748.98	9	83.22		
TOTAL	5024.15	19			

Having demonstrated that mean RTs for certain trials were different, the musical characteristics of these trials was examined. The shorter RTs for some trials indicated that those trials were generating clearer tonal implication, and were therefore abstracted more quickly to a tonal schema. For instance, trials in which the suffix note was included in the prefix generated shorter RTs. This is illustrated by the sixth experimental trial reproduced in Figure 6.3, where the D sharp suffix note is contained within the prefix.

Figure 6.4

Musical example of an experimental trial in which the suffix note is contained within the prefix

6 

6.3.2.5. Age group analysis of RTs to CORRECT trials

Since summation of means appeared to be disguising some of the variability within age groups, the means for each year group were examined to investigate possible differences between year–groups, and to determine whether the correlation between *same* and *different* trials was preserved. The resultant values are presented in Table 6.14. The inverse relationship between processing time and age is readily evident.

TABLE 6.14

Mean RTs of CORRECT responses
for each trial for each year group

Age group (n)	10–11 (5)	9–10 (7)	8–9 (5)	7–8 (9)	6–7 (7)
Trial					
1	134.5	147.4	147.2	143.4	236.5
2	140.2	143.9	174.0	151.6	157.7
3	152.4	145.2	161.0	182.1	263.8
4	150.3	149.8	195.3	138.5	172.0
5	135.4	238.5	163.5	161.9	150.7
6	118.3	140.1	131.4	120.2	150.0
7	131.7	130.2	134.5	122.8	134.6
8	143.2	146.0	189.0	187.9	189.3
9	165.6	132.0	186.3	163.7	175.2
10	143.2	130.3	162.3	167.9	164.2
11	139.8	218.5	162.0	158.8	162.0
12	118.5	154.5	171.3	197.6	289.7
13	137.0	155.0	146.3	300.3	164.3
14	128.6	157.3	172.0	159.2	188.5
15	164.2	191.8	165.0	172.2	187.7
16	141.5	118.1	112.5	139.8	223.5
17	142.3	133.2	143.0	140.7	148.0
18	168.3	191.2	136.0	141.8	195.5
19	127.0	130.7	129.4	163.4	158.3
20	135.0	139.7	122.3	167.9	128.5
Mean =	140.85	154.67	155.22	164.09	182.00
SD=	13.96	31.37	23.10	37.85	42.23

A one–way ANOVA of the mean values for each trial for each year group proved significant ($F(4,95) = 4.65$; $p = 0.0019$). The ANOVA Table is presented in Table 6.15.

TABLE 6.15

ANOVA of mean RTs of CORRECT responses
for each trial for each year group

ANALYSIS OF VARIANCE TABLE (ONE–WAY)

SOURCE	S.S	DF	MS	MSR	*p*
Years	18333.8	4	4583.45	4.65	0.0019
Residual	93645.8	95	985.75		
TOTAL	111979.6	99			

Differences in standard deviation between the 6–7 year olds (*i.e.* 42.23) and the 10–11 year olds (*i.e.* 13.96) proved significant confirming that older subjects respond more consistently.

The one–way ANOVA of the mean RTs of CORRECT responses for each trial for each year group therefore confirmed a significant difference between year groups in both experiments one and two, *i.e.* in both the non–contextualised and contextualised presentations of *same* and *different*.

A Tukey HSD test verified that the mean of 140.85 for the 10–11 year old group was significantly different (*i.e.* $p<0.05$) from the mean of 182 for the 6–7 year old group.[3] This confirmed the findings of the previous

[3] The post hoc Tukey Honestly Significant Difference test is highly conservative.

experiment.

Pearson product–moment correlation coefficients for CORRECT responses for all trials of different year groups proved non-significant, demonstrating the lack of correspondence between the age groups. The lack of correspondence between the summated response times and the individual year results demonstrates the variability of RT measures. One possible explanation for this, of course, is the fact that the representative year groups are very small in number (*i.e.* from 5 to 8 subjects per year) and insufficient responses are being utilised to calculate a mean of significant central tendency. It was therefore decided to assess test reliability by comparing repeated measures within the test.

6.3.2.6. Reliability of the measures

A measure of the reliability of the RTs could be provided by the agreement between the four *same* condition trials which were repeated in the experiment. This approach would correspond to the split–half method of reliability applied to same responses. A coefficient of agreement would constructively utilise the redundancy within the experimental design to produce an indication of the reliability of response times.

However, the RT scores of the 6–7 year olds were significantly different from those of the other age groups and were therefore removed from the test reliability analysis. RT means for each trial for the 7–11 age group only were computed. The mean RTs and standard deviations for CORRECT only responses to each experimental trial for the 7–11 age group (26 subjects) are given in Table 6.16. Bracketed figures show the means obtained from the 33 subjects from all years.

TABLE 6.16

Mean RTs and SD for CORRECT responses by the 7–11 age group (brackets give means for 6–11 age group)

Trial Number	Condition	Mean RT		Standard Deviation	
1	same	143.1	(162.4)	33.2	(93.4)
2	"	151.6	(152.8)	50.2	(46.1)
3	"	162.3	(185.7)	44.5	(121.7)
4	"	154.2	(157.8)	75.8	(68.6)
5	"	176.4	(171.1)	102.9	(92.4)
6	"	127.9	(132.3)	38.6	(42.1)
7	"	129.4	(130.3)	41.2	(37.7)
8	"	165.5	(171.0)	67.5	(70.5)
9	"	159.6	(162.9)	71.9	(64.8)
10	"	153.8	(156.1)	46.7	(48.2)
11	different	169.6	(168.7)	65.1	(61.2)
12	"	157.0	(174.3)	53.5	(105.1)
13	"	185.2	(181.0)	112.6	(101.2)
14	"	151.5	(155.2)	47.8	(47.6)
15	"	174.8	(176.7)	64.4	(60.4)
16	"	127.7	(143.0)	31.5	(61.3)
17	"	139.2	(140.5)	36.4	(36.2)
18	"	163.1	(166.7)	71.6	(69.8)
19	"	136.8	(139.4)	44.9	(42.8)
20	"	145.2	(143.8)	47.0	(45.2)

Mean (of CORRECT *Same* RTs) = 152.38 centi-seconds
SD (of CORRECT *Same* RTs) = 15.34 centi-seconds

Mean (of CORRECT *Different* RTs) = 155.01 centi-seconds
SD (of CORRECT *Different* RTs) = 18.32 centi-seconds

The data were consonant with the original analysis of the data from all years. The observed RTs of 152.38 and 155.01 were similar, although shorter than the corresponding mean RTs of 158.24 and 158.93 obtained from all the year groups. Standard deviations were also similar between the two analyses (*i.e.* 17.03 and 16.37 calculated from all the years and

15.34 and 18.32 with the RTs of year two subjects removed).

The two sets of RTs between conditions were still significantly positively correlated (r=+0.642; associated t value = 2.37; p= 0.044).

A two–way ANOVA (without replications) of the RTs of the 7–11 age group demonstrated that no significant difference existed between the conditions of *same* and *different* (F(1,9)=0.33; p=0.586, not significant). However, significant differences were still observed between the trials themselves (F(9,9)=4.43; p=0.019).

An examination of the variability of certain trials as revealed by the standard deviation demonstrates the reasoning behind the removal of the 6–7 age group RTs from the reliability analysis. For example, the large standard deviation of 121.7 for Trial 3 reduces to 44.5 with the removal of the 6–7 year olds RTs. The mean of 162.3 for the 7–11 age group consequently represents a closer estimate of the true population mean for the sample than the larger 185.7 for the 6–11 age group.

The data of the 7–11 age group was therefore utilised to obtain an estimate and indication of test reliability. Each higher response time was compared with the lower response time for the same four repeated trials as shown in Table 6.17.

TABLE 6.17

Mean RTs for CORRECT responses
to each repeated *same* experimental trial

Trial Numbers	Lower RT	Higher RT
2 and 3	151.6	162.3
4 and 5	154.2	176.4
6 and 7	127.9	129.4
9 and 8	159.6	165.5
Mean	= 148.3	158.4
SD	=14.0	20.3

Correlation between matched trials: r = +0.929

A positive correlation exists between these two set of scores (r = +0.929; associated t value = 3.54, $p<0.05$).[4] This significant correlation can be taken as a good indication of a coefficient of reliability, and confirms that differences are attributable to the effect of the musical characteristics of the trial materials on cognitive processes.

[4] A one–tail test of significance of the correlation being different from zero is appropriate since the arrangement of the four trials gives the higher comparison scores in one list.

6.4 Discussion

This experiment demonstrates that the notes of the prefixed tonic triad affected both error rates and response times.

6.4.1. Level of discrimination

The first null hypothesis (that there would be no significant difference in the distribution of the equiprobable classificatory responses of *same* and *different*) was rejected for all years except the youngest age group. Subjects were therefore able to clearly differentiate between the same and different contextualised paired stimuli.

6.4.2. Age and error responses

The second null hypothesis (that there would be no significant difference in mean error rates observed in different age ranges) was rejected. Subjects of different ages and different gender were clearly finding differential levels of difficulty with the trial materials. Moreover, developmental effects were confirmed by the significant correlation between decreasing errors with the increasing age of subjects.

6.4.3. RTs of *same* and *different* responses

The third null hypothesis (that the mean RTs of the conditions of *same* and *different* from all responses would not be significantly different) was not rejected from the analysis. This differs from the finding of the first experiment where a significant difference between *same* and *different* condition mean RTs was found. The addition of the context–defining tonal prefix presumably induced subjects to use similar processing strategies for both conditions.

The fourth null hypothesis (that mean RTs for the correct only responses would not be significantly different between the conditions of *same* or *different*) was not rejected. No significant difference was observed between RTs for *same* and *different* conditions, although there was a developmental effect apparent from the shorter RTs of the older subjects. Variance of RTs of the youngest age group were significantly different from other year groups. No significant difference in mean RTs for correct responses between the conditions of *same* and *different* suggests that the processing required for the mental translation of auditory stimuli into a verbal response is the same for both conditions.

The differences in variability observed both within and between the *same* and *different* conditions of experiment one were stabilised by the contextualised stimuli of experiment two. As the *different* condition variability of experiment two was less than that of the *same* condition, this suggests that subjects found the contextualised *different* stimuli of experiment two easier to discriminate than the uncontextualised *different* stimuli of experiment one. This supports the notion that abstraction to a tonal schema confers a processing advantage.

6.4.4. Differences between trials

The fifth hypothesis (that mean RTs for the CORRECT responses would not be significantly different between the trials themselves) was rejected. Unlike the previous experiment, the trials themselves exhibited significant differences (as demonstrated by the two–way ANOVA). Interestingly, the differences between *same* and *different* mean RTs observed in the previous non–contextualised presentation was negated by the context–defining prefix.

The results indicate that the children's recognition of semitone discrimination in both context–free and contextual presentations is progressively facilitated between the ages of six to eleven, with responses exhibiting fewer errors and decreasing reaction times with increasing age.

The observed significant difference between trials confirms that the responses were affected by the contextualisation and that reaction times indicate cognitive processing. Given the experimental design of repeated measures, systematic significant differences in reaction times observed in the contextualised presentation can be explained only by the effect of the triadic prefix affecting the abstraction to a cognitive set of perceptual hierarchical relations, or schema.

6.4.5. Systematic variability

The sixth null hypothesis (that the RTs should exhibit no significant degree of correlation between the matched trials of the two conditions) was rejected. Whereas no correlation had been observed in the non–contextualised presentation, the effect of the contextualisation of the triadic prefix on the responses was particularly marked for trial means. This correlation was observed for means calculated from correct responses from all subjects and correct responses from the 7–11 age group.

This significant correlation for contextualised comparisons confirms that the context–defining prefix was indeed affecting the paired suffix notes and has demonstrated that a triad is a sufficient context–defining stimulus for children. This confirms that the correlation of the mean reaction times of the matched trials in experiment two may be a function of cognitive abstraction to a tonal schema.

No significant positive correlation was observed between *same* and *different* semitones in context–free presentation in experiment one. The subsequent contextualisation in experiment two, induced by a triadic prefix to each of the notes forming the semitone, produced the significant correlation between the two conditions.

6.4.6. Processing strategies

It would seem from the results reported alone that two differing schemata were operating in this and the previous experiment. The previous experiment seemed concerned with perceptual matching, while this experiment seemed to indicate a schema that deals with deviations from a set of stimuli. The schema for the set is evoked by a group of pitches meeting some minimum criteria for associative behaviour or orientation. The contextual triadic prefix of this experiment serves this function. However, the absence of any context–defining prefix in the previous experiment may cause stimulus pitches to be perceived on a same/different protocol, probably requiring some form of retrograde processing from the second to the first stimulus.

The poor performance of the younger age children is consistent with Piaget's conception of pre–operational thinking. Operational thinking is characterised by the notion of reversibility in one of two forms: inversion (negation) and reciprocity. At the level of concrete operational thought, negation applies to classificatory operations, while reciprocity applies to operations involving relations. Younger children are likely to manifest behaviours demonstrating reversibility less successfully than older children, and so their poorer performance might be attributable to the inability to apply concrete operational thinking to the task.

6.4.7. Reliability

The redundancy of repeated trials within the experimental design (required to balance the number of trials within each condition of same and different) was useful in providing an indication of the consistency of the test. A significant degree of correlation was observed between trials with identical stimuli in the contextualised presentation, but absent from the non–contextualised presentation. This affirms the notion that the triadic prefix exerted an effect on the cognition of the suffix notes when contextualised within a tonal schema.

6.4.8. Conclusion

This second experiment is a novel demonstration of RT measures to produce a measure of the internalisation of pitches to auditory memory. It is proposed that the observed significant differences in RT responses serve as a measure of the internalisation of musical pitches to the cognitive structure of a tonal schema and that responses may therefore be classified according to a perceptual hierarchy. The hypothesis that the perceptual facilitation of the coding of redundancy within such a recognised and practised cognitive structure such as tonality is thus supported as a psychological reality for children of this age.

A further experiment was devised in order to gain further understanding of hierarchical cognitive processes The hypothesis which may explain the significant differences in RT concerns the extent to which cognitive abstraction is facilitated by the tonal strength or clarity of the stimulus. In other words, the tonal specificity of the stimuli was hypothesised as related to the observed RTs, with different RTs observed for those stimuli which were either greater or smaller in their tonal range of constituent pitches in relation to the circle of fifths. The further experiment designed to test this hypothesis is reported in the following chapter.

7. EXPERIMENT THREE: TONAL SPECIFICITY AND REACTION TIME: RESPONSES TO *SAME* OR *DIFFERENT* PAIRED-COMPARISON NOTES IN DIMINISHED TRIAD CONTEXT

7.1. Rationale

7.1.1. Introduction

The data obtained in experiment two revealed differences in RTs between the experimental trials and further evidence was required to explain these differences. Experiment three was designed to obtain data which would provide an indication of the musical features of stimuli that give rise to the differences. In particular, it was designed to determine whether tonal strength or specificity of stimuli could be related to the groupings of member notes around the circle of fifths.

This possibility is supported in the results obtained by Cuddy (1985), who found that a group of musically experienced listeners were able to differentiate sets of stimulus tones constructed to represent different degrees of spread around the circle. Her listeners judged stimuli on the criterion of appearing to '*go together as a group*' (Cuddy, 1985, p. 353), *i.e.* on the basis of this '*goodness of fit*'. A group of musically inexperienced listeners, however, failed to make these differentiations. These findings suggest that perception of tonality is related to musical experience or training.

Bharucha and Stoeckig (1986) also found evidence of differentiation by adult subjects on the basis of similar tonal groupings in their RT study of cognitive representation of harmonic organisation.

The ability of children to make discrimination on this basis appears not to have been investigated previously, as no *a priori* model for experimental procedure was found in the literature.

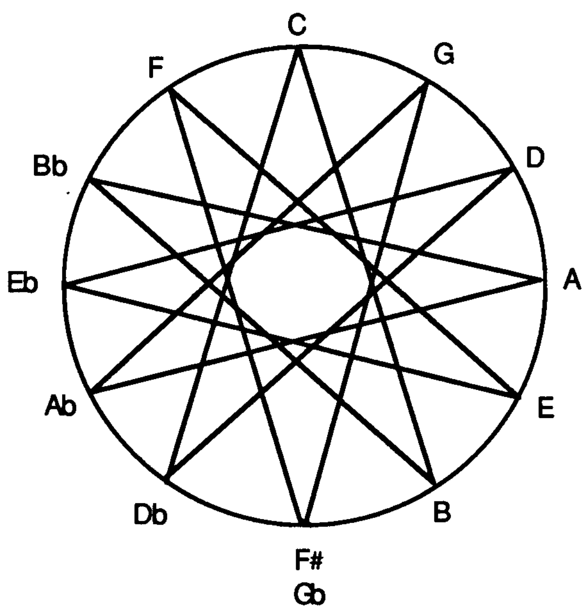
7.1.2. Influence of Previous Experiments

The validity and utility of chronometric measurement with children has already been demonstrated in experiments one and two. The purpose of experiment three was to answer some of the questions raised by the data of previous experiments. The relationship of keys to each other would be estimated by the degree of tonal spread between matched stimuli as indicated by the circle of fifths representation. The processing required for matching stimuli suggesting different tonal centres could invoke differing RTs. The *same* or *different* classification paradigm was used again to examine cognitive structures by comparing response times to near key relationship (*i.e.* proximal in the circle of fifths) with responses to far relations (*i.e.* non–proximal in the circle of fifths).

The polarisation of the implied tonalities of standard/comparison stimuli would be best achieved by tones occupying maximally distant positions on the circle of fifths. This influences how many notes a particular stimulus may have. A stimulus with five adjacent different notes constituting a minimal spread within the circle of fifths (*i.e.* a pentatonic set) does not provide the semitones of the major scale which are important indicators of tonality.

Figure 7.1

Semitone relations in the circle of fifths



The notes C, G, D, A, and E which are contiguous in the circle of fifths representation (Figure 7.1) are tonally ambiguous. They could belong to the major diatonic sets of F major, C major, or G major as well as their respective related keys of D minor, A minor, or E minor. Five–note pentatonic stimuli would not provide a clear enough tonal implication to investigate tonal specificity.

The full range of seven notes of the diatonic major scale is required to provide an unambiguous tonal centre. Presentation of stimuli comparing seven different notes might produce stimuli which are too long to recover accurate comparisons from younger children. However, it is not necessary to present all seven pitches to delimit a particular tonality. The tritone of C and F sharp/G flat suggests one of two possible tonal sets: either the G major set (*i.e.* the notes on the right hand side of the circle of fifths as presented in Figure 7.1) or the D flat major set (the notes on the left hand side of the circle of fifths as presented in Figure 7.1). The presentation of any other note with this tritone clearly specifies which tonal set is implied. Within the configuration as presented by Figure 7.1, any note on the left hand side of the circle of fifths implies the key of D flat major while any note on the right hand side of the circle of fifths would indicate the alternative key of G major. A context–defining stimulus can thus clearly be constructed to indicate a specific tonality by inclusion of a tritone and one other context defining note.

Certain configurations of tones are more tonally specific than others. *Tonal specificity* or *tonal spread* is determined by the number of adjacent pitches in the circle of fifths included in a stimulus. A stimulus may include notes which are members of a particular tonal set or may include tones which are outside this already indicated set of tones. For example, the notes C, B, F, and D are contained within only one specific major tonal set (*i.e.* C major). However, the notes C, B, F, and E flat are not members of any one tonal set. The notes C, B, F and D are contained within the *narrower* tonal spread of seven notes in the circle of fifths (i.e. F, C, G, D, A, E, and B), whereas the notes C, B, F and E flat are contained within the *wider* spread of eight pitches (i.e. B, G flat, D flat, A flat, E flat. B flat, F, and C).

An important consideration on the length of stimuli is its demand on memory in the mental abstraction required for comparison. The previous experiment involved relatively short four–note stimuli but abstraction to a tonal set might be facilitated if the stimuli were too long to be kept in short term auditory memory. Transference from short term auditory memory to a mental representation is required to ensure that the obtained RT is a measure of cognitive processing.

The cognitive abstraction of a musical stimulus might be facilitated by transposing the comparison stimulus. A transposition comparison does not allow note–for–note matching of absolute pitches, but rather it requires that the structural features of the stimuli have been understood and compared. A transposition task must involve a deeper level of abstraction than note–for–note matching of similar stimuli. Clearly, the task is different from those of the previous experiments, but the testing of the perceptual equivalence of transposed melodies would be an effective way of ensuring that assimilation to a scalar schema has ensued.

This raises a number of questions concerning tonality of the standard and comparison melodies. If the key of the standard stimulus is different for each trial, then this would mark an important change from the experiments reported in previous chapters. It would also be a different situation from those of probe–tone experiments where a repeated context-defining prefix continually re–establishes the same tonal context for each trial (*e.g.* Krumhansl, 1979). The changing of the tonal set for each trial can be avoided by some kind of fixed transpositions for comparison stimuli to produce a limited number of tonal sets.

Bartlett and Dowling (1980) examined the key–distance effect in a transposition task with adults (mostly musically experienced). The task involved the comparison of two melodies in which subjects had to detect whether the comparison stimulus was an exact transposition of the first using a four point confidence rating scale. The first experiment used two five–note melodies: three further experiments used familiar melodies. Six conditions were employed in the experiment, *viz.*

transpositions (T);
tonal lures in the same key as the standard (LS),
in a nearly related key (LN) or in a far key (LF);
atonal lures with the same contour as the standard (At);
and different contour tonal lures (D)

(Bartlett and Dowling, 1980, p. 504)

They found evidence to support a key distance effect in that items that were not transpositions but structurally changed to represent a far key (*i.e.* LF: tonal lures in a far key) were easier to reject than tonal lures in the same key as the standard (LS). Their fourth experiment looked at this key–distance effect with children (three groups with mean ages of 5.6, 6.9 and 8.6 years) with *same* and *different* responses. They found a key–distance effect but no other significant effects: this led them to the conclusion that with children of this age, the absolute pitch of notes is more salient than the interval between them. These researchers postulate that the key–distance effect is attributable to the assimilation of perceived pitches to a culture–specific musical schema which represents tonality, or sense of key. They conclude:

> *A critical test of pitch similarity versus a mode schema account might involve a rigorous control of the number of new pitches that distinguish LN and LF comparisons from standard stimuli. Then differences in false alarm rates to LN and LF comparisons having the same number of new pitches could not be attributable to absolute pitch memory and would support the schema view. Such an experiment remains to be done.*

(Bartlett and Dowling, 1980, p. 514)

Despite this evidence, a decision to avoid transpositions was made here on the basis that continual transposition between trials might hinder the establishment of tonal centres. The utilisation of near and far tonal relationships proposed in the *tonal specificity* experiment reported here is an alternative method of addressing the problem outlined by Bartlett and Dowling. Not only has the problem remained unexplored, but the RT methodology proposed is a novel and more precise indicator of cognitive processing than confidence ratings of *same* or *different*.

Croonen and Kop (1989) have followed up Bartlett and Dowling's experiments but the investigation has concerned the relationship between tonal information and interval information during specified retention time intervals (*i.e.* 1, 5, 8, 15, and 30 seconds). They used seven–note sequences with a clear tonal structure (*e.g.* C, E, G, F, D, B, C) and were attempting to examine the relationship between the tonal *clarity* or tonal strength of musical sequences and interval information. Their definition of *tonal clarity,* however, is simplistic and does not take account of the uniqueness of certain intervallic combinations.

> *The dimension of tonal clarity, as introduced, can be defined easily within the bounds of Western music. It is the degree to which a chord sequence establishes a particular key. As Schoenberg (1954/1969) states, there are three main triads, those at positions I, IV, and V. Traditionally, a "strong", or in our terms tonally clear, sequence is IV–V–I; all notes of the individual chords are diatonically related to a particular key, other possible keys (e.g. the dominant and subdominant) are implausible. In tonal music, chord sequences are possible that include tones that are far removed from the original key, as well as chords that do not uniquely point towards one tonal*

> *center. These structures are thought of as being* *<u>tonally unclear</u>.*
>
> *(Croonen and Kop, 1989, p. 64)*

This explanation of tonal clarity presents a number of difficulties for the musician. The classification of chords into those which are either *clear* or *unclear* in tonal implication is somewhat arbitrary. The ability of chords to specify particular tonalities is determined by the constituent intervals of those chords. A particular chord can be tonally unambiguous (*e.g.* a dominant seventh) or a chord can be ambiguous and specify a particular number of keys (*e.g.* the chord of C major might suggest the keys of F, C, and G). The notion of *tonal specificity* as defined in this study is a better descriptor of the dynamic nature of music. Although the chordal pattern of I–IV–V does specify a particular key in relation to functional diatonicism, individual triads themselves are poor specifiers of tonality as they do not include a tritone. The inclusion of notes which are far removed from the original key does not necessarily render them tonally unspecific: for instance, a modulation to a related key is extremely clear in tonal implication.

Bartlett and Dowling ignored the effects of contour in their conclusions. They used all possible contours of five–note and seven–note sequences in their experiments (Bartlett and Dowling, 1980). Edworthy's study (1985a) however, paid great attention to the effects of contour. The experimental sequences for experiment three were designed so that paired standard and comparison stimuli preserved similar contours to ensure that comparisons were not affected by differences of contour.

As the data collection process of the RT experiments of the study reported here involves individual testing, and is consequently time–consuming, a compromise between test duration and the number of possible conditions gave six possible conditions. The clearest tonal contrast is given by the two degrees of tonal specificity which polarise tonal relations between those of a particular tonal set (*i.e. narrow*) and those outside a particular tonal set (*i.e. wide*). As standard and comparison note groups can represent either of the two degrees of tonal spread, this gives a possible four combinations (*i.e.* narrow standard and narrow comparison, narrow standard and wide comparison, wide standard and narrow comparison, wide standard and wide comparison). Although this gives four *different* conditions, *same* stimuli can have notes which either belong or do not belong to the particular tonal set. This means that *same* stimuli can have only two conditions. The six conditions thus formulated provided the experimental design for experiment three as being likely to yield maximum information while still being practical with children.

The effect of rhythm was a final consideration. Experimenters have used either familiar sequences which have preserved the rhythm of established well–known melodies or they have used isorhythmic sequences devoid of rhythmical characteristics. Although the presence of a rhythmic identity may characterise stimuli as more musical, there is no doubt that alterations to the rhythm can affect the perceived identity of the sequence. This has been well demonstrated by Palmer and Krumhansl (1987a, 1987b). The use of isorhythmical sequences gives a better guarantee of freedom from intrusive variables.

7.1.3. Hypothesis and Statistical Decisions

The experimental hypothesis was that:

No significant difference will be observed in reaction times or error rates in respect of *same* and *different* paired–comparison stimuli where these belong to closely related tonal schemata (proximal in relation to the circle of fifths) or are distantly related.

The data recorded by the computer system consisted of the dependent variables of the RT taken to respond to a trial and the category of response. Analysis was abbreviated from that used in the previous experiment as developmental factors could be revealed by inclusion of the age groups as one of the factors of a factorial analysis. The experiment examined not only the between–groups factor of age as an independent variable but possible differences attributable to the within–group factors of trial type (*i.e. same* or *different*) and key relation (*i.e. narrow* or *wide*). This generated a three factor design. A three–way factorial design ANOVA was used to estimate the significance of the variables.

A further intention was to quantify a hierarchy of perceptual relations, or at least provide some evidence concerning the relation of keys to the circle of fifths in children's cognitive processing of music. An individual analysis of trials, rather than an individual subject analysis, was more likely to reveal the extent to which matched trials (grouped on the basis of specified criteria) were related perceptually. One of the problems with subject analysis noted in the previous experiments was that wide subject differences had been observed. The computation of mean RTs for each trial condition could ensure that this subject variability would not bias any of the mean RTs. It would not matter if a particular subject was generally slow to respond if that slowness was a component of each obtained mean. This would be important for data which shows large differences of subject variability. A possible method of removing this subject variability would be to use a base–line RT as a measure of covariance (*e.g.* Fiske, 1982a; discussed in chapter five).

7.1.4. Experimental Design

In the previous experiments reported here, problems with analysis by age group were experienced when the number of available correct responses on which to base a true estimate of RT was insufficient. Separate analysis of year groups ideally requires a larger number of responses from subjects than previously obtained, perhaps with some replications to ensure that the obtained mean responses are a clear indication of intentions. Multiple responses may be a good solution to this problem, particularly as the correlation between different blocks could be computed to confirm the the responses were truly indicative. Furthermore, the training effect of the experimental procedure is likely to be minimised with multiple responses from fewer subjects. The pilot studies reported in chapter four showed that the first few experimental trials produced slower responses until subjects became familiar with the procedure. Responses from a larger group of subjects are likely to maximise this experimental training effect. Some experimenters have favoured multiple responses with fewer subjects: for example, Edworthy (1985a) used only ten subjects for her experiment.

The subjects for this third experiment were taken from the same school as previous experiments as the children already had experience of the experimental procedure. This avoided the need for children who were unfamiliar with the experimental procedure to reach a training criterion before testing. Practice trials were considered undesirable as they would be time–consuming and affect the length of the experiment, at risk of inducing test fatigue. The use of subjects already experienced with the experimental situation avoided such pretesting and ensured more realistic response time measurements. The number of trials utilised by Edworthy (1985a) was 2240, though these were spread across 14 conditions (two tasks with seven melody lengths) for her ten subjects with 16 melody pairs. The proposed six conditions of this experiment would require fewer responses to achieve representative mean RTs for each condition.

The design was again conceived as a computer–controlled forced–choice binary *same* or *different* response to comparison stimuli which were classified into one of six conditions employing various combinations of tonal relationships according the proximity of tonal relation of the final comparison note of each standard or comparison stimulus. Two types of relation were defined, *i.e.* stimuli with all notes within a specified seven–note tonal span (*i.e. narrow*) and stimuli with a wider span than seven notes of a chromatic scale (*i.e. wide*). It was hypothesised that the tonal spread (*i.e.* the range of notes around the circle of fifths) of the comparison stimuli when contextualised by the diminished triad would influence the RT to respond to the stimuli.

The presentation order of the trials was randomised to avoid the effects of test fatigue, particularly with younger subjects.

7.2. Method

7.2.1. Subjects

Sixteen children aged 7–11 from the rural primary school in Derbyshire used in the previous experiment acted as subjects. Two boys and two girls were randomly chosen from each of the four year–groups. All the subjects (except one) had undertaken the second experiment the previous year and were therefore familiar with the procedure and method.[1]

7.2.2. Materials, apparatus and procedure

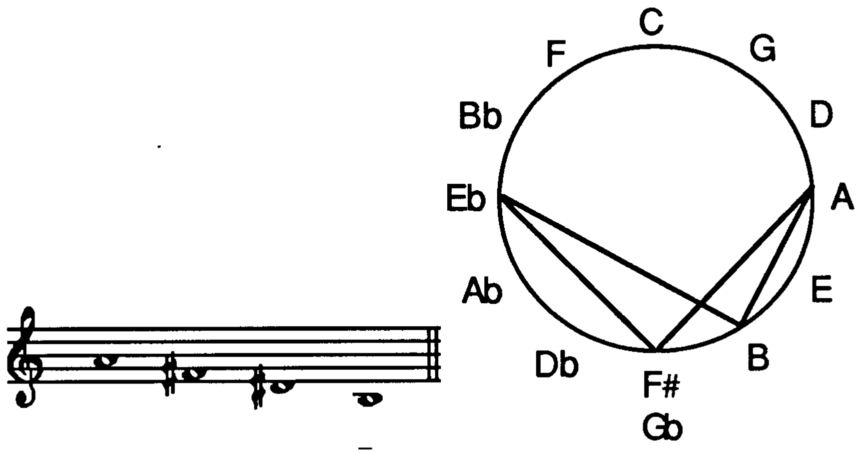
Each test item comprised a standard and a comparison stimulus pair, each made up of a sequence of four tones. Stimulus pairs were either the *same* or *different*. The first three tones (*i.e.* prefix) of each four–tone sequence were always the same [A, F sharp, and D sharp]. The final tone (*i.e.* suffix) was varied so as to create two experimental conditions: *narrow* and *wide*.[2] Suffix notes ranged over a major seventh from B flat to A. This range was chosen as the range likely to be within the singing compass of the subjects (cf. Welch, 1979, Welch *et al.*, 1989).

[1] The inexperienced subject was one of the older pupils and found no difficulty with either the experimental situation or procedure. No practice trials were considered appropriate.

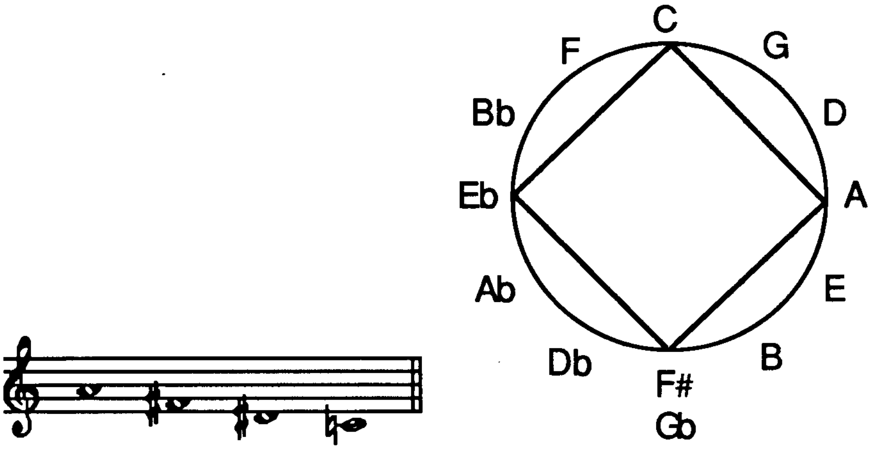
[2] The terms *narrow* and *wide* were considered to best describe the suffix and prefix. Other possible descriptors (*e.g. near* and *distant, proximal* and *non–proximal, member* and *non–member)* were rejected as they did not describe the complete stimulus, but explained only the relation of the suffix to the prefix.

The tonal structure of a four–note sequence was designated as:

Narrow: If the suffix tone belonged to the tonality established by its preceding three–note sequence (i.e. if the spread of the four tones around the circle of fifths did not exceed 7 adjacent positions), *e.g.*



Wide: If the suffix tone did not belong to the tonality established by its preceding three–note sequence (i.e. if the spread of the four tones around the circle of fifths exceeded 7 adjacent positions), *e.g.*



The possible combinations of *narrow* and *wide* sequences within the two categories of *same* and *different* stimulus pairs thus created six possible experimental conditions:

Condition	Standard/Comparison	Standard	Comparison
1	same	narrow	narrow
2	same	wide	wide
3	different	narrow	narrow
4	different	narrow	wide
5	different	wide	narrow
6	different	wide	wide

There were six trials for each condition, producing 12 *same* trials across two conditions and 24 *different* trials across four conditions.[3] An example of an experimental trial for each of the six conditions of various degrees of tonal spread or *tonal specificity* is given in Figure 7.2.

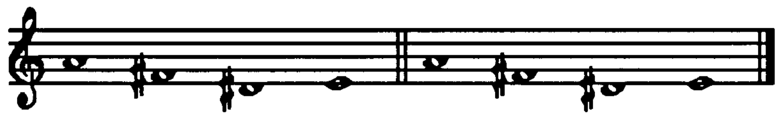
[3] A musical representation of the thirty–six trials of the experiment is given in Appendix VIII and the computer program is reproduced in Appendix IX.

Figure 7.2

Musical example of each of the six experimental conditions

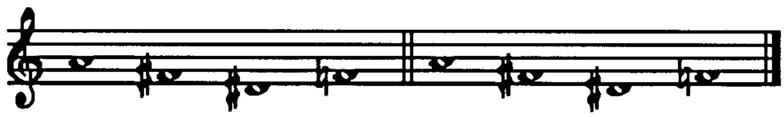
Condition 1: Same Condition, Narrow Standard, Narrow Comparison

Trial 1 (SNN)



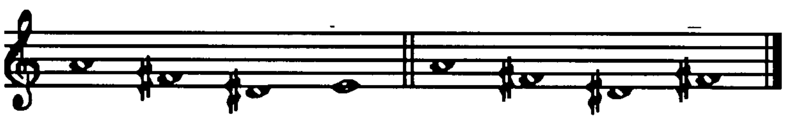
Condition 2: Same Condition, Wide Standard, Wide Comparison

Trial 19 (SWW)



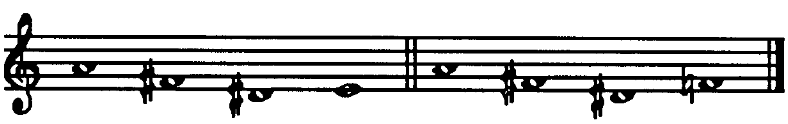
Condition 3:Different Condition,Narrow Standard,Narrow Comparison

Trial 2 (DNN)



Condition 4: Different Condition , Narrow Standard, Wide Comparison

Trial 3 (DNW)

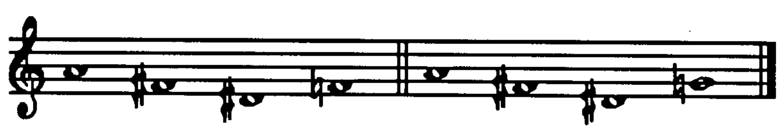


Condition 5: Different Condition, Wide Standard, Narrow Comparison

Trial 20 (DWN)

Condition 6: Different Condition, Wide Standard, Wide Comparison

Trial 21 (DWW)



7.3. Results

7.3.1. Levels of performance

The mean and standard deviation for the number of CORRECT responses by subjects for each trial condition is given in Table 7.1. The maximum number of correct responses per trial was sixteen. Error rates ranged from 8% for trials in the first condition (SNN) to 31% for trials in the fourth condition (DNW).

TABLE 7.1

Mean and standard deviation for the number of CORRECT responses to trials in each condition

Condition		Mean Correct	% Incorrect	SD
1	(SNN)	14.67	8%	1.03
2	(SWW)	13.33	17%	2.07
3	(DNN)	13.17	18%	1.17
4	(DNW)	11.00	31%	3.35
5	(DWN)	11.17	30%	2.86
6	(DWW)	13.67	15%	1.21

(N=16)

(S=Same; D=Different; N=Narrow; W=Wide)

It can be seen that those paired stimuli with a mixture of *narrow* and *wide* suffix notes induced more errors since they had a lower mean correct rate and greater variability of responses and correspondingly greater standard deviations. A one–way ANOVA of the number of correct responses for each trial yielded significant differences in the mean error rates between the six conditions ($F(5,30)=2.76$, $p=0.0363$).[4]

A significant negative correlation was found between the number of correct responses and the mean RTs for all trials ($r=-0.506$, $p<0.01$).[5]

7.3.2. Reaction Times

The mean correct RTs for the thirty–six trials were normally distributed within the range 117.83 centi–seconds to 204.14 centi–seconds with one outlying trial mean of 240.14 centi–seconds. The mean of the thirty–six trials was 158.6 and the standard deviation was 26.15 centi–seconds.

The mean and standard deviation for the mean correct RT for each of the six experimental conditions is shown in Table 7.2

[4] Nonparametric Kruskal–Wallis one–way ANOVA was also significant ($\chi^2=11.66$, $p=0.0397$)

[5] Since the first three conditions had almost identical means (*i.e.* 153.95, 153.12, and 154.78), Pearson's *r* was used rather than ranking as it is more sensitive to the size of differences between scores.

TABLE 7.2

Mean and standard deviation for the mean CORRECT RT responses (in centi–seconds) for each condition

Condition		Mean	Standard Deviation
1	(SNN)	153.95	12.13
2	(SWW)	153.12	25.33
3	(DNN)	154.78	17.54
4	(DNW)	162.83	19.72
5	(DWN)	191.05	34.10
6	(DWW)	135.86	12.33

(S=Same; D=Different; N=Narrow; W=Wide)

A oneway ANOVA of the mean correct RTs of each condition proved significant differences between the experimental conditions (Table 7.3).[6]

[6] Nonparametric Kruskal–Wallis oneway ANOVA was also significant (χ^2=12.78, p=0.0256).

TABLE 7.3

A one–way ANOVA of the mean RTs for CORRECT responses for each condition

ANALYSIS OF VARIANCE TABLE (ONE–WAY)

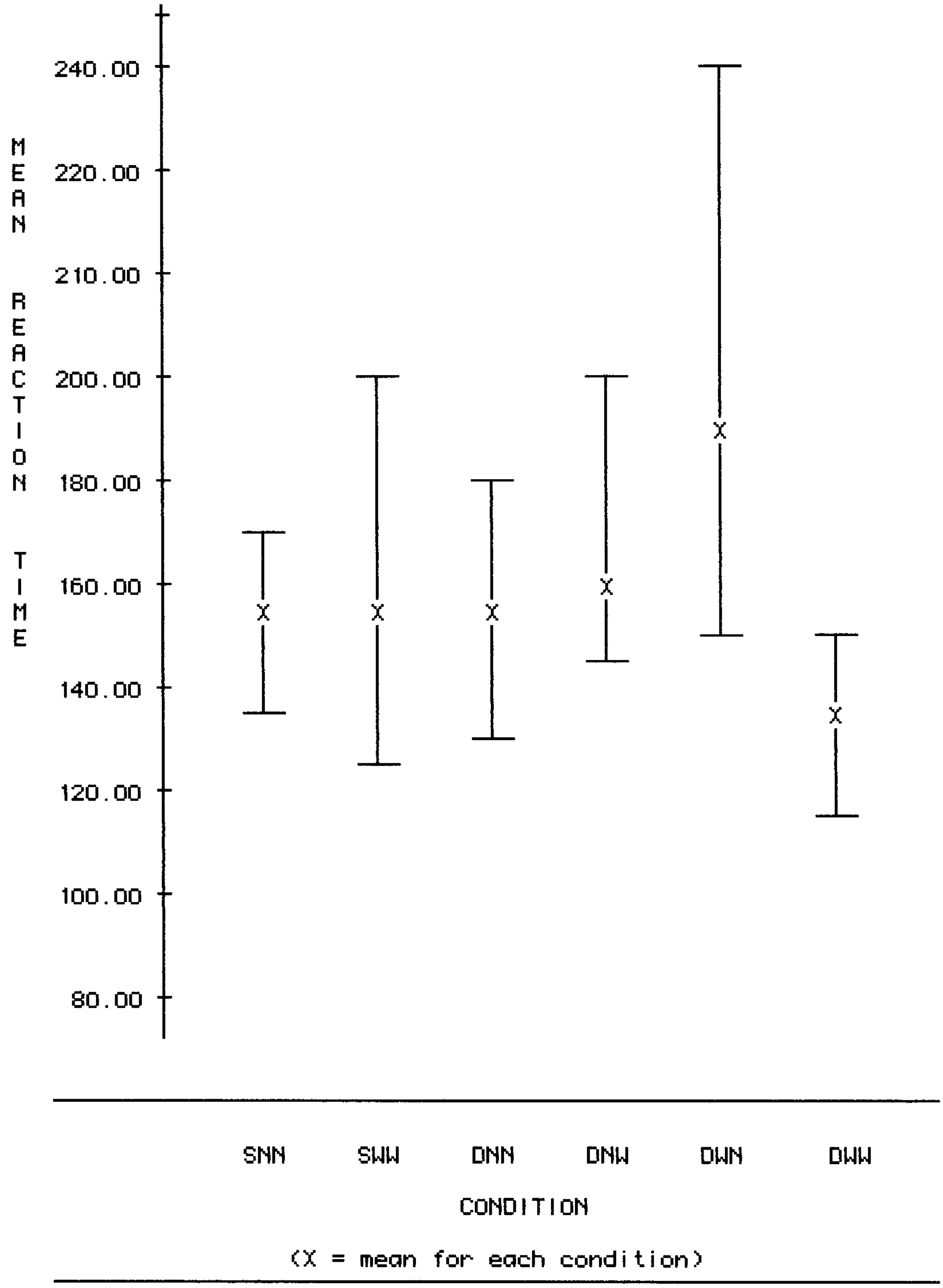
SOURCE	S.S	DF	MS	MSR	*p*
Between Conditions	9926.75	5	1985.35	4.25	0.0048
Within Conditions	14001.34	30	466.71		
TOTAL	23928.09	35			

A *post hoc* Tukey HSD (Honestly Significant Difference) test found that condition 5 was significantly different from condition 6 at the $p<0.05$ level.

A graphic comparison of the ranges of the mean correct RT responses (in centi–seconds) shows clearly the relationship between the *same* and *different* conditions and is presented in Figure 7.3.

Figure 7.3

Ranges of the mean correct RTs (in centi-seconds) for the six conditions

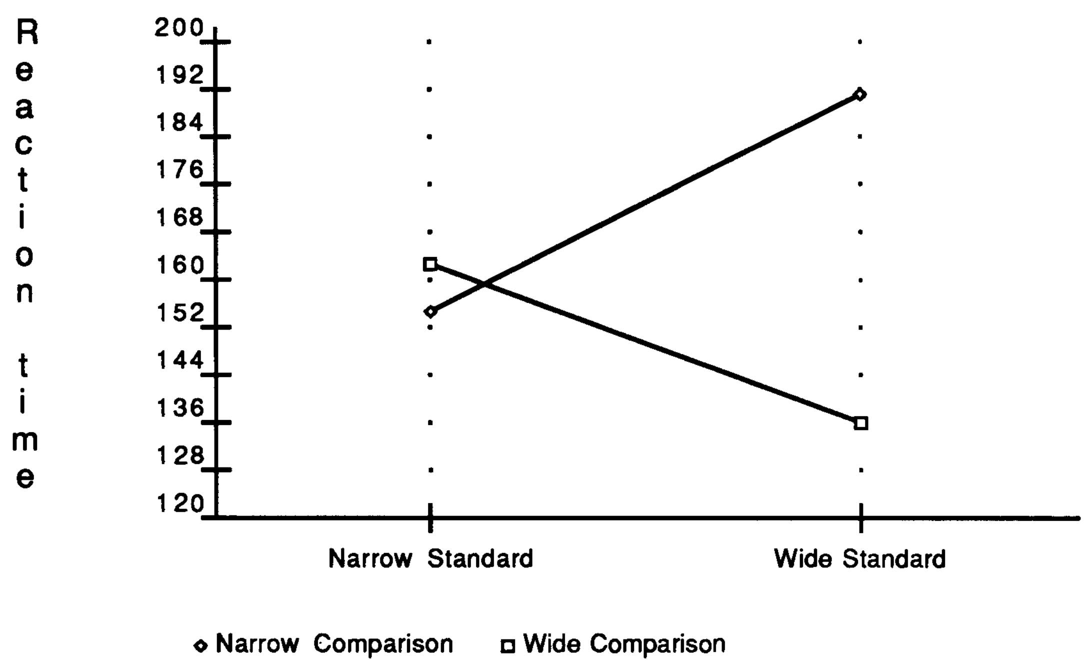


This graphic comparison of mean RTs for each condition shows clearly the differences in mean RT and differences in variability, particularly for the fifth and sixth conditions. What is also apparent from this comparison is that the *wide* condition suffix standard note does not produce the same effect under the two conditions of DWN and DWW. What is remarkable is that the fastest mean response time for the DWN condition (*i.e.* 150.07 centi–seconds) is slower than the slowest response for the DWW condition (*i.e.* 149 centi–seconds).

A two–factor within–subject MANOVA of the *narrow* and *wide* conditions, partitioned into the two factors of the suffix notes of the *standard stimuli* and *comparison stimuli*, revealed that these main factors demonstrated a significant interaction, although neither of the individual factors themselves were significant. Figure 7.4 shows the relationship between the mean RTs for the standard and comparison stimuli notes for each different condition.

Figure 7.4

<u>Relationship of mean RTs (in centi-seconds) of the *standard* and *comparison* suffix notes for each *different* condition</u>



A significant interaction indicates that the effect of one of the variables is not the same under all conditions of the other variable, hence the non–parallel lines. This interaction indicates that the effect of the *narrow* and *wide* suffix notes were not the same under the various experimental conditions. This is particularly noticeable for the *wide* standard note where *wide* comparisons generally generated faster responses than *narrow* comparisons. Since subject differences were equally distributed across the mean RTs for each test item, the interaction must be attributable to trial differences rather than a response characteristic for each subject. The MANOVA table is presented in Table 7.4

TABLE 7.4

MANOVA of the mean RTs for within subject factors of *narrow* and *wide* factors for each *different* condition

MULTIVARIATE ANALYSIS OF VARIANCE TABLE

SOURCE	S.S	DF	MS	MSR	*p*
Between Conditions					
Constant	623105.82	1	623105.82	3337.08	0.000 **
within cells	933.61	5	186.72		
Within Conditions					
Standard	129.60	1	129.60	0.34	0.585
within cells	1901.73	5	380.35		
Comparison	3333.03.	1	3333.03	4.19	0.096
within cells	3973.21	5	794.64		
Interaction	6000.53	1	6000.53	9.24	0.029 *
within cells	3247.73	5	649.55		

(* $p<0.05$; ** $p<0.0005$)

7.3.3. Subject differences

Subject differences were investigated by computing response time means for correct responses for each subject for each condition. Means for the six conditions revealed a similar profile to the trial means and a MANOVA investigating within subject effects confirmed the result of a significant interaction ($F(15,1)=9.71$, $p=0.007$) between the suffix notes of the standard and comparison stimuli.

However, an ANOVA of the mean RTs of the six conditions by the age group of the children revealed no significant differences between the age groups.

7.4. Discussion

7.4.1. Error rates

The significant differences in the mean error rates for each trial between the conditions confirmed that the manipulation of the experimental materials was responsible for the differences. Those paired stimuli which mixed the *narrow* and *wide* tonal spread in relation to the circle of fifths produced more errors than the other conditions and subjects were much less consistent in their responses. This effect can be explained as a result of the cognitive confusion which results from the priming effect of the initial stimulus. If the standard stimulus (of the diminished triad and context defining note) suggests a different tonal set from the comparison stimulus, this may prove more difficult for the brain to interpret. This is consistent with the notion that a tonal schema is not invoked by a tonally inconsistent matching stimulus pair.

5.4.2. Error responses and RTs

The significant negative correlation which was found between the number of correct responses and the mean trial RTs suggests that the observed differences between the conditions are systematic and that the error rates and RT responses reflect cognitive processing strategies. Similar systematic difference was observed in the previous experiment where the context was generated by a diatonic triad. The diminished triad context seems be partially responsible, in conjunction with suffix notes, for the systematic differences between the conditions and this suggests that the diminished triad possesses tonality–defining properties for children. This is in marked contrast to the experiment of Cuddy and Badertscher (1987) who found that the diminished triad did not generate a context sufficient to recover the tonal hierarchy with children.

7.4.3. Mixed standard–comparison conditions

Although the RT trial means were normally distributed between a minimum of 118 centi–seconds and maximum of 204 centi–seconds, trial 26 (condition DWN) possessed an outlying mean score of 240 centi–seconds. The condition of this outlying trial *(i.e.* different: *wide* standard – *narrow* comparison) has larger mean RTs for correct responses coupled with greater variability of both RT and error responses. These attributes suggest that children found this condition most difficult to classify correctly. This difficulty with classification can be explained as a result of the effect of priming of the other conditions. Reaction times to *correct* responses were shorter and more consistent in those conditions in which the standard context defining stimulus generated a narrow tonally specific context. However, condition six (different: *wide* standard – *wide* comparison) produced consistent shorter and more accurate responses than any of the other different conditions.

Certain trends are observable between those *same* and *different* conditions in which the comparison stimulus represents a different relation in the circle of fifths from the standard, *i.e.* DNW and DWN. In mixed conditions (*i.e. narrow* and *wide*, *wide* and *narrow*), RTs are generally slower and possess more variability, hence the greater standard deviations than other conditions. There is greater variability in those stimuli which do not relate to a specific tonality.

7.4.4. Identical standard–comparison conditions

Although the means are almost identical (*i.e.* 153.95 and 153.12) and similar to those of the previous experiments, the standard deviation of the SNN condition is 12.13 whereas the standard deviation of the SWW condition is 25.33. These unequal distributions suggest that different processing strategies are being employed. The RTs for the DWW condition are generally faster. This difference is important in that it seems to suggest that pitches are abstracted to a tonal schema, but that if the stimulus has a greater tonal spread than that found in a specific single tonality the rejection of *same* can be made more quickly than if both sets of notes suggest the same tonality and have to be compared at a lower level in the perceptual hierarchy.

This faster processing of stimuli which occurs when both standard and comparison notes fall outside the tonality suggested by the context–defining prefix is suggested by the shorter RTs of the sixth condition (DWW). The significant difference between this condition and the others is obvious as the correct mean RTs for the two conditions do not overlap at all (Figure 7.3).

This finding is inconsistent with the findings of Janata and Reisberg (1988), reported in chapter two. They hypothesised that notes that were more stable within a given context would be more quickly recognised as belonging to that context. The shortest mean RT obtained in the third experiment reported here was for *different* stimuli in which standard and comparison notes were *wide* (*i.e.* DWW).

The DNN condition RT mean is very similar to the SNN condition RT mean. The similarity in processing times of the DNN condition to those of the SNN condition suggests that pitches are initially abstracted to a tonal schema before an alternative comparison strategy determines whether there is a difference between the two stimuli. The slightly longer time for the DNW condition might be attributable to the priming effect of the standard suffix note suggesting a particular tonal schema followed by a comparison outside the tonal schema primed by the standard stimulus. This seems to work in reverse for the DWN condition. The initial stimulus does not unequivocally specify a particular tonality so the cognitive processing mechanism is not primed. When the second comparison stimulus does specify a tonality, it takes significantly longer to reach a classification decision. Abstraction to a tonal schema ensues before the stimuli can be classified as different.

However, when both suffix notes fall outside a clearly defined tonality, abstraction to a tonal schema does not occur for the comparison pitches and they are more quickly and consistently classified as different since the suffix note is more easily remembered. Other factors such as pitch height may also be important for this contextual comparison.

7.4.5. Conclusion

Reaction times were longer and errors more numerous for the two conditions which involved a mixture of *narrow* and *wide* tonal spread. This suggests that tonal specificity of the stimulus systematically affects the responses. The significant ANOVA and multiple comparison Tukey test highlights that the relationship between the factors here is a complex one. The significant interaction between the standard and comparison suffix notes of the stimuli confirms this priming effect of the standard suffix note.

These differences relate to the way in which the information is processed by the brain. The semantic memory non–redundant hierarchical model (Collins and Quillian, 1969; *cf.* Figure 2.3), discussed in chapter two, suggests that each aspect of categorical information is stored at the highest appropriate level within the hierarchy. As each cognitive concept is linked to only one other concept, retrieval involves a spread of activation from each unit to adjacent nodes. Those comparisons which are closer together take less time to process. This is supported by the findings of this experiment, as RTs were longer and errors more numerous for the two conditions which involved both *narrow* and *wide* tonal spread (*i.e.* DNW and DWN) than for conditions which presented the same degree of tonal spread (*i.e.* SNN, SWW, DNN and DWW).

The priming effect of the standard stimulus is influencing the speed of response for those conditions in which the comparison stimulus contradicts the tonal expectation generated by the standard stimulus. The particularly greater mean RT for the condition with different tonally *wide* standard and tonally *narrow* comparison suggests that the cognitive mechanism is not primed by the *wide* standard stimulus, but that the tonal implication of the *narrow* comparison stimulus is processed before a response is made.

Although consistent subject differences were observed in the previous experiments, no significant difference was observed between age groups with this randomised small subject sample. Some of the older subjects, for example, were relatively slow to respond and often inaccurate while some of the younger subjects were quick and accurate. This subject variability was examined by an additional experiment reported in the following section.

The RT responses indicate that the notion of tonal specificity of a musical stimulus affects the time taken to respond to the classification task of the experiment. The greater variability of the *different* responses observed in the previous experiment is here clarified in that the variability is at least partially attributable to the tonal specificity of the stimuli.

7.5 Supplementary investigation of subject differences

7.5.1 Rationale

The differences in RT observed between the conditions could not be attributed to subject differences as the means obtained were computed by summing across all subjects for each trial. This effectively disguised the wide differences in RT observed between subjects in all the experiments reported here.

An investigation of the reasons for this wide subject variability examined two possible causes. One possible cause could have been differences in the cognitive abilities of the subjects. It had been possible at the time of testing to ascertain from the class–teacher the reading ages of the children as classified by the Suffolk Reading Scale (1981). The relationship between cognitive ability (as shown by the Reading Test) and RT could thus be explored.

Another possible cause could have been differences in motor–response time taken by subjects in converting cognitive decision to an operational response. The subjects of this third experiment were further tested to obtain a measure of psychomotor response time. A computer program measured the time each subject took to press the space bar on the computer keyboard to each of 20 randomly presented beeps.[7] The delay between beeps varied between three and eight seconds. As mentioned

[7] The computer program is given in Appendix X.

previously, the use of responses to random beeps has also been used by Fiske (1982a), who used the responses as a covariate in a factorial experiment.

A beep (bell–sound on the computer) was used for this supplementary investigation. If a musical note on an electronic keyboard had been used, part of the delay could be attributed to the durational properties of the stimulus itself. A beep was considered appropriate as it has a minimal duration. Furthermore, a beep would help to negate any response delay attributable to the rhythmic character of a multi-note sequence. RTs may be affected by the speed of presentation of tones in a multi-note sequence, although this was not tested in the experiments reported here. The space bar was considered appropriate for responses as a binary choice response was not needed and it seemed unnecessary to complicate the response mechanism with the button box. A choice decision would have slowed down the response time.

As two subjects of the third experimental subject group were not available at the time of testing, only 14 mean baseline RTs to the beeps were obtained.

7.5.2 Results

Mean response times for each subject were calculated across the six conditions of the third experiment. The distribution of RTs to the third experiment for the fourteen subjects who participated in this additional experiment had an overall mean of 153.7 with a standard deviation of 60.3.

As the error rate was the other dependent variable in the previous experiments, the relationship between errors and other variables was investigated. The mean number of correct responses to the 36 trials was found to be 29.1, with a standard deviation of 5.4.

The reading ages as obtained on the Suffolk Reading Scale for the fourteen subjects gave a mean reading age of 108 with a standard deviation of 12.8.

Although there was a positive correlation between the reading ages as obtained on the Suffolk Reading Scale and the mean response times for the musical stimuli for each subject, the correlation was low and not significant (r=0.211, NS).

The number of correct responses in experiment three was also moderately positively correlated with reading ages, and this relationship proved significant at the 5 percent level for a one–tailed test

(r(12 *d.f.*)=0.462, p>0.05). The number of correct responses was also positively correlated with the mean RTs for the musical stimuli, although this was not significant (r=0.297, NS). The trade–off between the speed and accuracy of response in previous experiments was not preserved in this experiment

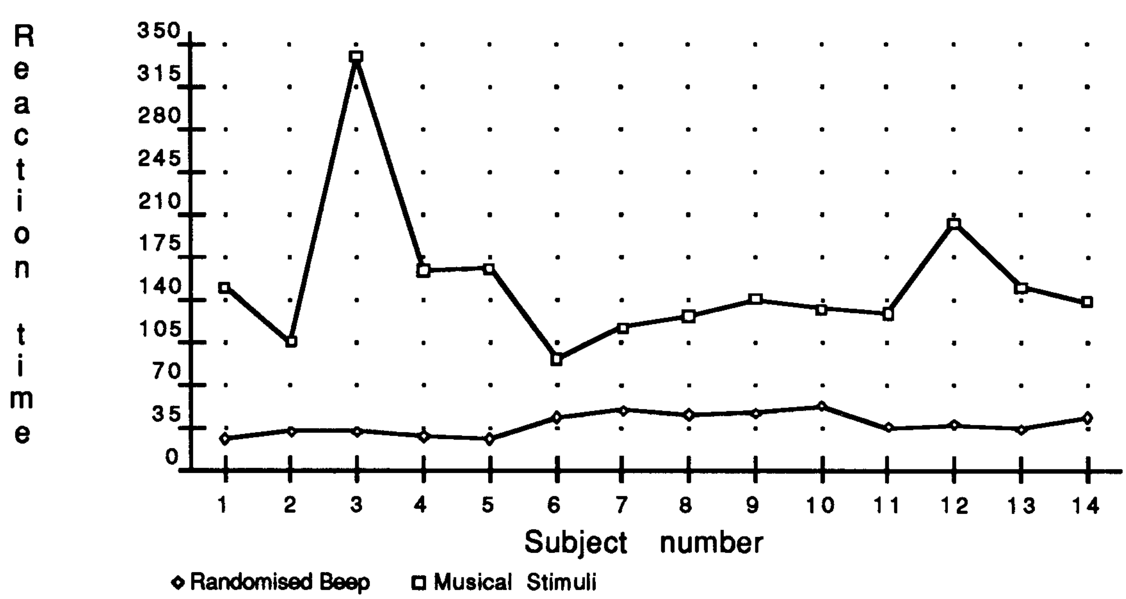
The mean of psychomotor response times to the randomised beeps was 37.6 centi–seconds. This time is much shorter than the mean RT taken to respond to the musical stimuli (*i.e.* 153.7 centi–seconds). There was also much less variability between subjects in their responses as the standard deviation of RT to randomised beeps was only 9.1 centi–seconds as compared with the 60.3 centi–seconds for musical stimuli.

The correlation between RTs for responses to randomised beeps and those relating to musical stimuli was non–significantly negatively correlated (r=–0.039). This low correlation indicates the absence of a linear relationship between the two RT scores and precludes the analysis of RT scores with the use of randomised beep scores as a covariate. A condition of covariate analysis is a significant correlation between the covariate and other dependent variable (*cf.* Bryman and Cramer, 1990). This proved not to be the case here. Furthermore, since the beep responses are not systematically related to the the musical stimuli responses, there is no need for covariate analysis as the error variance does not have the experimental effect masked by psychomotor response time.

However, the negative correlation coefficient does not reveal clearly the relationship between the variables as they are not equally distributed as revealed by the different standard deviations. The response times for randomised beeps all lie close to the minimum time presumably needed for a psychomotor response: the response times for musical stimuli, on the other hand, exhibit much greater subject variability. This is shown clearly in the graphical representation of the relationship between the response times for both the musical stimuli and randomised beeps for the fourteen subjects (Figure 7.5).

Figure 7.5

Relationship between the response times (in centi-seconds) for both the musical stimuli and randomised beeps



This diagram makes clear that a great deal of subject variability cannot be attributed solely to differences in response mechanisms.

Reading test scores and RT to randomised beeps were negatively correlated (r=–0.402).[8] This negative correlation, although not significant, is consistent with some level of relationship between RT measures and reading test scores. Those subjects who responded faster to the beeps generally had higher reading ages.

7.5.3 Discussion

The lack of significant correlation between the Reading Scale and the mean response times for the musical stimuli for each subject suggests that the RT responses are not linearly related to the cognitive abilities as measured by the reading test.

However, the significant correlation between the number of correct responses and reading ages suggests the presence of a relationship. Those children with higher reading ages were consistently achieving higher numbers of correct responses, although in some cases this adversely affects the RT score. For example, subject three scored high on the reading test (*i.e.* 119), was highly accurate (*i.e.* 35 correct responses out of 36 responses), but took much longer than other subjects

[8] This failed to reach the critical value for significance at the 5 percent level for a one–tailed test (i.e. r(12 d.f.)=0.458).

to make the response (*i.e.* mean response time of 340 centi–seconds). This is an example of the speed–accuracy trade–off discussed in chapter two, and indicative of personal response style of individuals. Subjects who wish to maintain accurate responses invariably take longer in response. This was confirmed by the positive correlation between the mean RTs for the musical stimuli and the number of correct responses, although this correlation was not significant.

The clear difference between the mean RTs for responses to musical stimuli and randomised beeps demonstrates that the significant differences in the RT for musical stimuli is not determined solely as a function of psychomotor response. This affirms that observed differences are a result of cognitive processes. The variability, therefore, can be attributed to differences in strategies adopted by subjects in their responses, with a speed–accuracy trade–off creating subject variability.

The negative correlation between the Reading Test and the randomised beeps, although not significant, is in accord with some level of relationship between general cognitive ability and speed of response. Those subjects who responded faster to the beeps had higher reader ages.

The interrelationship of these variables explains some of the wide subject differences observed. While it is recognised that not too much weight should be attributed to these other variables, they help to explain the complexity of the cognitive mechanism governing response time, and particularly the way in which the processing strategy adopted can affect the observed processing time. Such variables are useful for indicating where further research could be particularly helpful in clarifying the role of other cognitive factors in RT studies.

This brief supplementary investigation indicates that general cognitive ability, as measured by a reading test, is related to the cognitive processes which govern music perception as measured by RTs or the number of correct responses to a choice RT paradigm. However, psychomotor response times are not responsible for the observed differences in either RT or error rates. This is further evidence to support the psychological reality of cognitive music processing as demonstrated in this third experiment.

CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

8.1 Summary of experimental results

These experiments have examined the cognitive structures that children employ when listening to musical pitches. They have demonstrated the utility and validity of RT measures to identify perceptually salient factors in musical pitch perception. Some of these factors have not previously received attention in the research literature, and the results therefore present some new information.

The pilot experiment contrasted three types of stimuli and significant differences in variability of RTs for both children and adults demonstrated that the method had general applicability for investigation of musical cognition. The suggestion that the processing of musical pitch invokes cognitive structures which are related to processing time was supported by this experiment.

Children between 8 and 9 years old in the pilot experiment discriminated the interval of semitone between matched short three–note stimuli at a level beyond chance, suggesting different degrees of tonal spread in relation to the circle of fifths. The longer and more variable RTs for processing *same* condition items support an exhaustive processing model.

Cognitive models formulated by such investigators as Bamber (1969) to explain *same* response superiority seem unnecessarily complex for the musical comparison task of the pilot experiments. Certainly, auditory tasks are distinct from visual tasks and no *identity–reporter* type mechanism of the type proposed by Bamber is needed to interpret the results. The fact that response times are subject to experimental manipulation, as explored in chapter two, might be one possible reason why RTs for *same* were longer: the instruction to subjects was to respond as soon as they were sure of making the correct response. This cautious judgment equates to the manipulation of RT by verbal instructions reported by Ratcliff and Hacker (1981). The finding of longer response times for *same only when sure* is consistent with processing along more dimensions for the *same* condition.

The first experiment, although simplistic in requiring comparison of uncontextualised notes, was helpful in clarifying a number of important issues. Most important was that the experimental method itself was not responsible for the differences in RTs observed in the pilot. If processing of *same* and *different* stimuli leads to significant differences in RT, this could be considered to be an artifact of the experimental situation. This experiment showed no such difference. This therefore confirms that the significant differences in RTs observed in all of the other experiments are attributable to cognitive processes rather than an aspect of the experimental situation.

A clear relationship between decreasing RTs and increasing age emerged in this first experiment. A significant correlation suggests that children between the ages of six to eleven find increasing ease in discriminating consistent with other research using RT responses. The correlation was significant for children in the six to eleven age group. This correlation was also observed in the second experiment.

Although no significant difference was observed in mean correct RTs between uncontextualised *same* and *different* conditions, significant differences in RTs were observed between these conditions when suffix notes were each contextualised by a major triad prefix. Furthermore, while no significant positive correlation was observed between *same* and *different* notes in context–free presentation, the subsequent contextualisation by a triadic prefix to each comparison suffix–note produced a significant correlation. This is a particularly important finding, because it suggests that the contextualisation effects are systematic. This is evidence of time for cognitive processing and the fact that shorter RTs were observed for certain stimuli suggests that higher–order cognitive functioning results from contextualised stimuli.

The third experiment using a diminished triad prefix confirmed that the tonal specificity of stimuli was related to the RTs of responses. Significant differences in RTs of correct responses were found for those stimuli which were different in their tonal range of constituent pitches in relation to the circle of fifths. However, this result is inconsistent with the research of Cuddy and Badertscher (1987) who found that the diminished triad was insufficient to generate a tonal hierarchy. It is clear from the experiment reported here that the diminished triad can be a sufficient context-defining stimulus capable of generating a tonal context. This may be because only one type of prefix was used within any one of the experiments reported here, whereas three context types were used by Cuddy and Badertscher. Furthermore, the repetition of the tonic note within the prefix might be responsible for part of the experimental effect observed by Cuddy and Badertscher and an effect of short term memory processes.

A particularly interesting finding of the third experiment was the relationship between *wide* and *narrow* stimuli defined by the specified pitches within a particular tonal set. The significant difference in RTs between those *different* condition stimuli which both began with notes outside the implied tonal centre of the experiment (*i.e.* E major) are consistent with abstraction to a tonal schema. The finding that comparison of suffix notes outside the tonal schema suggested by the prefix could result in significantly faster RTs was unexpected and inconsistent with hypothesis of Janata and Reisberg (1988). However, it was consistent with the notion that children must be abstracting the stimuli to a tonal schema. An *identity-reporter* processing stage between more extensive serial searches might explain the faster RT for those stimuli which began with notes not as easily abstracted to a tonal schema. The experiments reported here indicate that this holistic processing to a tonal schema tries to make sense of the stimuli before the serial search takes place. This accords with research in other domains. For instance, Ratcliff (1981) has observed that RTs from the comparison of strings of five letters indicates comparison by the amount of overlap between stimuli, and not letter by letter comparison. The shorter RTs for the DWW condition suggest that non-instantiation of a tonal schema confers a processing advantage for the comparison of tonally–inconsistent suffix tones. This initial processing stage is consistent with the model proposed by Fiske (1990), described in chapter two.

The tension between psychoacousticians and musicians, as outlined in the first chapter, in identifying the problem of musically valid simulation against controllable experimental material is clarified by these experiments. A four–note stimulus is indeed sufficient to generate context–defining musical materials and consequently constitutes a musical environment capable of invoking a set of tonal relations.

It is proposed that the observed differences in RT responses serve as a measure of the internalisation of musical pitches to a cognitive structure such as a tonal schema and that responses may therefore be classified according to a perceptual hierarchy of pitch relationships. The hypothesis that perceptual facilitation of the coding of redundancy within such a recognised and practised cognitive structure such as tonality is supported for children aged between seven and eleven by these experiments.

8.2 Possible developments of chronometric methodology

The experiments here have established the utility of the chronometric technique to investigate perceptual and cognitive processes. However, the binary choice RT paradigm is not the only method available for RT measurement. It would be possible to increase the number of choices to a stimulus, operating in the manner of ratings scales, which have proved useful in providing more precise indicators of judgments than a two choice response. A task involving more than two response categories could be expected to increase the time to undertake the response. More than two possible responses have been used by experimenters: for example, Nettelbeck and Brewer (1976) used an eight choice task in which stimulus lights were either close (directly above) or distant (i.e. 2.8 metres) from the response keys.

A greater number of response categories than two could be applied to RT experiments with music stimuli and might give greater dispersion of response times than would the simpler tasks of the experiment reported here. The strength or certainty of the response could also be a factor in the decision–making in multiple response tasks, and this is not always directly related to the RT. For instance, an indecisive response is not measured by a binary choice paradigm. The *not sure* classification is not accounted for by a forced choice binary response. In fact, a long RT for a particular response, *same* or *different*, might not indicate a preference for a particular response, but indicate an indecisive guess response by the subject. It is unfortunate that such longer RTs generated by guessing have a much greater effect on mean scores, and particularly on variance, than is desirable. This has been observed in some of the experiments reported here and was a reason for some particularly large scores being discarded from analysis.

Multiple response classification would be a useful extension of these experiments and task strategy could be controlled by such experimental manipulation. The investigation of subject differences in the third experiment suggests that different strategies are being used by individual subjects. The third subject, for instance, took far longer than others to respond, but maintained an almost entirely correct response pattern as a consequence. The inclusion of additional response categories could serve to highlight differences in RT responses, particularly if this were linked to the confidence of the response. This is directly comparable to the classification task of of *goodness of fit* response as used by Krumhansl. For example, subjects could be asked to indicate a *yes/no* response on a five–point differential scale.[1] A tentative hypothesis would be that more easily classified items would be responded to not only more quickly but to a different response category. This could help to polarise differences in RT responses. Such a confidence rating has been used by Bharucha and Pryor (1986), who used a four point rating scale with confident and not–so–confident responses for both *same* and *different* conditions.

The technique of chronometric measurement of responses could also be applied to other types of test where multiple responses are required. For

[1] The BBC micro analogue port would allow interfacing with a multiple response button–box.

example, a pitch discrimination test which required one of five notes to be identified as different could be linked to a multiple response button–box where the buttons are positionally linked to specified notes and the task involves identification of the position of the altered note. This approach could identify recency effects in that the position of the altered note in the pitch task (or rhythm task) could be manipulated to investigate whether an effect of serial search is perceptually salient. This feature has deliberately not been investigated by the experiments in the study reported here as the *standard* and *comparison* notes were always in terminal positions to avoid serial effects. However, analysis of serial placement effects would provide further information concerned with either perceptual or cognitive processing.

The technique developed here would allow a variety of musical processes to be investigated. One such example is the study of perceptual streaming (Bregman, 1978; Bregman and Campbell, 1971; Van Noorden, 1975, 1982). For instance, the perceptual effect of two–part polyphonic music could be explored by having two buttons designated as either *upper* or *lower* part respectively. Subjects could be asked to press the appropriate button to discriminate a specified change in either of the parts and the RT to respond measured. Perceptual processes would be readily probed by modifications to the chronometric analysis approach. Many existing classification techniques (e.g. *semantic differential* and *personal construct theory*) could be enhanced by complementary engagement of RT measurement.

The wide subject differences revealed in the experiments reported here deserve further investigation. The relationship of musical discrimination ability to other cognitive abilities could be fruitfully explored. The factor of intelligence as revealed by a reading test revealed that there is a relationship between the correct classification of responses and intelligence. This factor of cognitive ability does not explain all the differences observed in the experiments reported here, and more research is needed.

The differences in the proportions of correct classification of responses of experiment three also support the conclusion that different strategies are being used by subjects in their judgments. The adoption of particular strategies may be related to other cognitive abilities. In fact, it is highly unlikely that music cognition is independent of other perceptual and cognitive processes. Further research is needed to clarify the relationship between music cognition and other cognitive abilities.

The relationship of psychomotor processes to chronometric measurement of this type is another factor which affects the speed of response to response time measurements of the type used in the study here. The third experiment suggested that the differences observed between subjects in RTs to musical stimuli were not affected by general psychomotor response mechanisms (*cf.* section 7.5). The effect of this factor needs to be identified in the interpretation of the speed of cognitive processes.

Another factor which merits further research is the role of cerebral dominance. Although both ears were used in this experiment, stimuli presented to different ears could help locate possible hemisphere dominance. Pechstedt *et al.* (1989) have reported evidence that tonality processing is located in the left hemisphere. RT measures could help confirm or refute this finding.

The findings of experiment three suggest that hemisphere specialisation is not responsible for the observed differences in RT. As the brain is contralaterally linked to the nervous system, superiority for one of the conditions of *same* or *different* could indicate that one side of the brain is more concerned with the processes involved in the discrimination of stimuli. However, significant differences were observed between conditions of the *different* response which were all generated by the right hand in experiment three. This could be consistent with a superiority for tonality processing in the left hemisphere, but this requires further research with dichotic presentation of stimuli.

The present study has not investigated the possible hierarchical structure of perceptual and cognitive domains. The differential RTs of discrimination tasks could be used to indicate the relative importance of processing mechanisms in cognition. For instance, as a simple discrimination task is likely to take less time than a more difficult task, subtraction of RTs could give an indication of processing time. This could have been applied to the differential times observed in the third experiment where the difference between mean response to musical stimuli (*i.e.* 153.7 centi-seconds) and randomised beeps (37.6 centi–seconds) is over 1 second (*i.e.* 116.1 centi–seconds). However, the rhythmical structure of the time-based presentation of the stimuli could be responsible for some of this one second delay: it is unlikely that cognitive processing takes all of this time.

Another aspect which could be explored is the quantification of tonal specificity as defined in chapter three. Two different maximally different relationships were explored in the third experiment. However, it is possible to categorise the intervals of the the major scale into seven levels of tonal specificity. The tritone is most specific and the octave or unison the least. This requires further experimentation, and would provide additional evidence for the reality of intervallic rivalry as a perceptual mechanism.

The use of more multiple response categories might be explored to see if greater RT is required for more complex tasks, and the understanding of what constitutes a more complex task could use the subtractive method developed by Sternberg (1969) to reveal if a hierarchy of cognitive relations is psychologically relevant.

The effect of training is another variable which could receive attention in extensions of the studies reported here. Morrongiello (1992), in her review of the effects of training on children's perception of music, argues that training serves to facilitate the speed at which interval and contour information is encoded (Morrongiello, 1992, *p*. 38). The closed environment of computer–presented musical stimuli with feedback mechanisms as part of the feedback loop could investigate whether such recognition is enhanced by practice. The trainability of pitch discrimination could be systematically explored. The effects of positive reinforcement and negative feedback, as well as no feedback whatsoever, could explore the ability of subjects to improve discrimination. A computer system allows the development of a complex training environment as well as an instructional mode of learning. A computer system can allow two–way interchange between subject and experimental environment and it is possible to use a pitch–to–midi converter to transform vocal responses to a particular pitch which could be interpreted by the computer. This methodology could investigate pitch predominance in the presentation of musical stimuli. For example, Temko (1971) asked subjects to sing the predominant pitch they perceived after hearing various extracts from twentieth century music. A computer environment could enhance Temko's methodology in allowing a two–way interchange between subject and experimental situation.

The use of RT measures could explore the cognition of more complex stimuli than simple melodic note-strings. The rhythmic and harmonic aspects of musical materials could be utilised as test materials. Furthermore, the influence of other variables such as age, gender and musical background could be observed. The experiments here have identified that these variables are important in cognition, but further work remains to be done to elucidate more fully the cognitive processes involved in musical cognition.

Alternative analytical methods than those used in the studies here may lead to important methodologically independent corroborative proof of models of cognition obtained by other experimenters. For example, RT measures might be used to generate a proximity matrix to show relationships from which hierarchical clustering which could be used to reveal relationships. Moreover, multidimensional scaling from the proximity matrix would also reveal relationships between the variables. The recovery of models by other experimenters such as Shepard's cyclical model of pitch (*cf.* Figure 1.1) or Krumhansl's conical configuration of musical pitch (*cf.* Figure 1.2) would further develop the psychological reality of the models.

8.3. Implications for further research

The perceptual relevance of the *tonal hierarchy* theory has been applied to children's cognition in only a limited number of studies (*e.g.* Krumhansl and Keil, 1982; Cuddy and Badertscher, 1987). The use of RT measures to demonstrate the tonal hierarchy has been applied in only one experiment, that of Janata and Reisberg (1988; reported in chapter two). The more recently postulated *intervallic rivalry* theory (Butler, 1989) has not formerly been applied to children's cognition.

These two models, although different, are not mutually exclusive. Pitches are related to an imagined tonic in both theories, but candidates for this tonic can change as a particular musical stimulus progresses. Thus the schema is dynamic with both theoretical notions, although the nature of the transformation is different. The tonal hierarchy theory assumes that pitches are related to an abstracted reference point and that the relationship of pitches is static and fixed. Krumhansl's conical projection of tonal relationships is invariant. However, music is not as fixed as this theory would appear to suggest. No account is taken of large-scale cognitive structures. Cook (1987b) has argued that the notion of large scale *tonal closure* has perceptual relevance for music listeners, although he found that tonal closure had a relatively weak influence on the perceptions of listeners. He found the effect of tonal closure limited to fairly short time spans. The large scale tonal implications of diatonic compositions can be accounted for by Schenkerian notions of dynamic tonal structure, however, and attempts have been made to relate functional harmonic theory to music cognition. Bharucha (1984), for example, has examined the directional properties of non–essential notes in what he terms *anchoring effects* in the resolution of dissonance

According to the intervallic rivalry theory, cognitive processing determines candidates for the tonic according to the best information available to the cognitive framework. Furthermore, processing acknowledges the influence of the local context, the temporal relations of notes in actual musical usage which are meaningful musically. The tonality of a musical phrase can be determined from its constituent intervals by intervallic rivalry processing, and the total absence of a tonic within a musical phrase is still capable of providing an unequivocal tonal centre. For example, although the key centre suggested by a diminished seventh chord is clear, the key centre changes without being established in the Wagnerian usage of two juxtaposed diminished sevenths with chromatic appoggiaturas. The key centre is not fully established as there is no perfect cadence, and the implied tonics may even be missing from the music. Music like this is not well explained by the static tonal hierarchy theory. Krumhansl has admitted the limitation of her theory:

> *Certain general cognitive–perceptual principles and capacities may also be found to emerge that apply to music of other cultures and more recent styles of Western music, whose theoretical description is at present less well–developed.*
>
> *(Krumhansl, 1983, p. 60)*

The theoretical description of other types and styles of music is not less well developed although it might be less well understood by psychologists.

The nature of the intervallic rivalry theory takes better account of the dynamic nature of the passage of music. The intervallic rivalry theory has been expressed simply as:

> *Any tone will suffice as a perceptual anchor – a tonal center – until a better candidate defeats it.*
>
> *(Butler, 1989, p. 238)*

This theory has been supported by the experiments here, especially the third experiment which used a diminished triad to generate a tonal context, something not yet replicated by tonal hierarchy experiments. The use of RT measures to explore these theories is novel. The accurate measurement of cognitive processing time is a quite different technique from the rating judgments required by probe tone methodology.

The experiments reported here provide a novel demonstration of children's cognition in that RT measures have not been used by other researchers to explore children's pitch processing strategies. The technique might be extended and used with younger children than those in this study (*i.e.* six year olds). Children as young as six were able to significantly differentiate the interval of a semitone in the study reported here. The age at which children begin to make this distinction has not been determined from this study. The discrimination task between *same* and *different* is simple but the concepts are still independent. A better methodology with younger children would require children to answer *yes* or *no* to the question which asks if the two things being compared are the same. This is an easier task as only one response classification is required and consequently attention to one concept instead of two. Such a response classification could influence the discrimination strategy used by subjects, and could affect RT responses. Subjects might be encouraged to adopt a *safe same strategy* which would favour longer *yes* responses to accommodate the rechecking necessary to establish a *same* response. This has not been tested in the present study.

Some interesting features of children's musical development might be answered by extensions of the chronometric procedure developed in the experiments reported here. For example, the distinction between training and development is something that could be investigated. The trainability of pitch discrimination is something not systematically investigated with young children and how this affects greater musical development in production tasks like singing or playing a musical instrument. The experimental system devised here could be programmed to provide an interactive environment for training in pitch (or rhythm) discrimination. The development of training materials with the computer process allowing feedback is an environment not yet exploited with children.

The experiments reported here have resulted in the development of research materials and techniques applicable to adults as well as children. A development of the materials need only change the *data* statements within the computer program to produce an environment which can inform psychologists of adult's processing mechanisms. Groups of subjects whose cognitive processing is being investigated (*e.g.* blind musicians with absolute pitch) can still be investigated to see if different processing strategies are utilised.

Computer modelling of the cognition of musical pitches has begun to receive attention recently. Simon (1968), with his LISTENER program determined an algorithm for finding the tonality of a given piece by frequency counts of the notes within a composition. However, this does not take account of the recent developments in computer programming with computers that simulate human behaviour in learning how to do a task. Recent developments employ neural networking, where links between nodes of related concepts are strengthened by positive reinforcement.

Bharucha (1987, 1991) has also formulated approaches to the tonality induction problem. He has attempted to demonstrate that a nodal network is capable of simulating cognitive behaviour and proposed a connectionist model. Bharucha has formulated a self-organising network model of musical harmony, MUSACT, which organises pitches according to chords and keys. Bharucha and Olney (1989) further discuss the relationship between tonal cognition, artificial intelligence and neural nets.

Scarborough *et al.* (1989) see such a network in terms of pitch class nodes which are linked, mirroring the perceived similarities of notes, chords and keys in line with the model of tonal pitch space proposed by Lerdahl (1988). Scarborough *et al.* (1991) outline the advantages and disadvantages of the linear tonal induction network which they have devised. They relate their network to the goodness–of–fit probe tone ratings obtained by Krumhansl.

Krumhansl (1990a) found that the that frequency counts of the pitches in a composition can give an indication of the tonal centre of a tonal piece: by weighting each note according to its duration, she was able to identify tonalities of the major key preludes of Bach's *Well–tempered Clavier.* Using only the first four notes of all voices of each prelude, she was able to find the key for 44 of the 48 preludes (*i.e.* 92 per cent accuracy). She attaches a great importance to the statistical distribution of notes and the consequent implied tonality:

> *...humans (as well as other organisms) are highly sensitive to information about frequency of occurrence. Thus, the primary significance of the observed correspondence between statistics of music and psychological data in these cases is to suggest a mechanism through which principles of musical organisation are learned.*
>
> *(Krumhansl, 1990a, p. 315)*

Scarborough *et al.* (1991) applied a frequency occurrence approach to their network of pitch–class nodes linked to key nodes, and found that they were able to identify accurately the key for all 24 preludes from Book One. They point out the problems with this approach to frequency note counts. Their linear tonality induction network includes the intervals of the third and fifth in the key and chord nodal connections, but the model does not contain the intervals of the minor second and tritone which are most tonally specific according to the *tonal rivalry theory* postulated by Butler (1989). Furthermore, the network does not take account of the order of presentation of pitches, or of any implied or perceived metrical structure of the musical stimuli. Brown (1988) has clearly demonstrated the importance of the presentation order of pitches in tonality perception.

The programming of neural networks and the simulation of cognitive behaviour is complex and the ideas of *heuristic* programming have not yet been explored in tonality induction.. As an adjective, heuristic means *'serving to find out or to simulate investigation'*.[2] Computer Scientists offer a different meaning for the term, for example Feigenbaum and Feldman (1963) have explained the use of the term as follows:

> *A heuristic (heuristic rule, heuristic method) is a rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces. Heuristics do not guarantee optimal solutions; in fact, they do not guarantee any solutions at all; all that can be said for a useful heuristic is that it offers solutions which are good enough most of the time.*
>
> *(from Firebaugh, 1989, p. 105)*

Heuristics, therefore, are rules of thumb used to provide a quick solutions to problems. Heuristic programming differs from algorithmic programming in that a heuristic does not necessarily provide a solution to a problem, unlike an algorithm. If the intervallic theory of pitch relationships mirrors how we ascertain the tonality of particular sequences of music, then this can be applied to analysing music. As certain intervals are more tonally specific than others (*e.g.* the tritone and

[2] Collins English Dictionary

minor second) then a computation of these intervals with weightings generated in a neural network simulation might better model the cognitive processes of the mind. The development of such a networking system for learning to recognise different styles by the speed and movement of tonal centres would be particularly useful. It is possible that such a system could recognise atonal music by the very absence of an established tonal centre.

Research in computer modelling of musical pitch has demonstrated the importance of the circle of fifths representation. Leman (1991a,1991b) has applied the neural network technique based on self–organisation developed by Kohonen (1984), the *Kohonen Feature Map (KFM)*, to investigate tonal relations. His neural network of 400 neurons was trained with 115 different chords (major and minor triads and seventh chords including augmented and diminished permutations). The neuron in the network that responds most strongly to a given input is called the *characteristic neuron* for that input. The neurons surrounding this characteristic neuron which are activated more highly than the remaining of the network are known collectively as the *response region*. Leman recovered a configuration of *characteristic neurons* loosely conforming to the circle of fifths in his experiments and argued that aspects of tonality can be explained by internal representations that develop through self–organisation. Furthermore, he has developed this notion of overlapping *response regions* as analogous to chordal facilitation and tonal functioning.

Leman's work is exciting in addressing issues of cognitive functioning in tonal organisation, particularly in recovering the configuration of the circle of fifths, but it could be argued that the network seems more concerned with the acoustical properties of the training materials rather than the functional diatonicism implied by chord progressions, although Leman does distinguish between *sensory* and *cultural* bases of his model. This is reflected in some of his descriptions of chords, *e.g.* his so–called *C dominant seventh chord* is really the dominant seventh of F major. The chords used by Leman to train the network include the interval of the minor second and tritone, unlike the linear tonality induction network of Scarborough *et al.* (1991). The KFM network developed by Leman is therefore more representative of chordal configurations than the network developed by Scarborough which does not use the interval of either the minor second or the tritone.

A model of artificial intelligence for music tonal induction would be very useful in education. Computer programs are far from intelligent in converting a music performance into music notation. The idea of an intelligent computer sequencer which is able to automatically interpret a tonality and provide a key signature for the music is attractive and would be very useful in educational contexts. At present, users of computer music systems have to specify such details as clefs, key signatures and time–signatures to display a notated composition with any accuracy.

A further application of a tonal induction computer simulation relates to artificially intelligent music performance systems currently being developed. Clarke (1989) is developing a system in the LISP environment which stores human performance data which can be used to model artificial musical performances. Music performance of tonal music is dependent on an understanding of the structural characteristics of the music. For example, phrasing in tonal music is closely linked to the harmonic implications of the music, which demands an implicit knowledge of functional diatonicism on the part of the performer. The performance of a chromatic appoggiatura in a tonal melody, for instance, is likely to dictate implied phrasing to a performer. Artificial music performance systems need to take account of the relationship between such structural characteristics and expressive musical performance and the tonal induction heuristic could serve as an important part of a cognitive modelling system.

While it is attractive to speculate about cognitive modelling as developing the understanding of perceptual processes, this experiment suggests particular pedagogical approaches to music teaching and learning. If abstraction to a scalar or tonal schema is developed first then the training of children's aural abilities might be best accomplished by scalic passages and discrimination of general intervallic types within a scale rather than particular types of interval (*i.e.* major, minor, diminished, augmented). This equates to Balzano's RT finding of scale step equivalence.

Finally, although this study has not concerned itself with musical meaning, some speculation of the role tonal expectancy plays in giving meaning to music might be considered. (Swanwick, 1973) found that an ability to make predictions concerning what followed a pair of alternating notes repeated as a basic norm seemed crucial to the understanding of music. There can be little doubt that the mind does make predictions about the future course of events in music, and that abstraction to a tonal schema is important for aesthetic meaning. The conflict between *schematic* and *veridical* expectancy, as defined by Bharucha and Todd (1989), where schematic expectancy is related to culture–based generalised expectancies and veridical expectancy is specific to a particular instance, is at the heart of all music meaning. The tension between what is expected and what is actually heard (*cf.* Meyer, 1956) is partly dependent on tonal expectancy and schema representation has meaning within a specific culture. The study reported here demonstrates that children make use of such cognitive structures in their music listening behaviour.

REFERENCES

Agmon, E. (1989). A mathematical model of the diatonic system. *Journal of Music Theory*, 33, 1–25.

Bachem, A. (1950). Tone height and tone chroma as two different pitch qualities. *Acta Psychologica*, 7, 80–88.

Bachem, A. (1954). Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 26, 751–753.

Badertscher, B. (1985). *Representation of the tonal hierarchy in children and musically untrained adults.* Unpublished honor's thesis, Queen's University, Kingston, Ontario.

Balzano, G. J. (1977). Chronometric studies of the musical interval sense. Unpublished doctoral dissertation, Stanford University. *Dissertation Abstracts International*, 38, 2898B.

Balzano, G. J. (1980). The group–theoretic description of twelvefold and microtonal pitch systems. *Computer Music Journal*, 4, 66–84.

Balzano, G. J. (1982). Musical versus psychoacoustical variables and their influence on the perception of musical intervals. *Bulletin of the Council for Research in Music Education*, 70, 1–11.

Bamber, D. (1969). Reaction times and error rates for 'same'–'different' judgments of multidimensional stimuli. *Perception & Psychophysics*, 6, 169–174.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* London: Cambridge University Press.

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key–distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 501–515.

Bartlett, J. C., & Dowling, W. J. (1988). Scale structure and similarity of melodies. *Music Perception*, 5, 285–314.

Berry, W. (1976). *Structural functions in music*. Englewood Cliffs, New Jersey: Prentice– Hall.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, 16, 485–518.

Bharucha, J. J. (1987). Music cognition and perceptual falicitation: A connectionist framework. *Music Perception*, 5, 1–30.

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological perspective. In P. M. Todd & D. G. Loy (Eds.), *Music and Connectionism* (*pp*. 84-127). Cambridge, Massachusetts, MIT Press.

Bharucha, J. J., & Olney, K. L. (1989). Tonal cognition, artificial intelligence and neural nets. *Contemporary Music Review*, Special issue on Music and the Cognitive Sciences, S. McAdams (Ed.), Vol. 4, 341–356.

Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies as a function of context. *Cognition*, 13, 63–102.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics*, 40, 137–141.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 403–410.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics*, 41, 519–524.

Bharucha, J. J., & Todd, P. M. (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal*, 13(4), 44–53.

Bisanz, J., Danner, F., & Resnick, L. B. (1979). Changes with age in measures of processing efficiency. *Child Development*, 50, 132–141.

Boltz, M. (1989). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4), 749–761.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, 18, 389–431.

Bregman, A. S. (1978). Auditory streaming: Competition among alternative organisations. *Perception & Psychophysics*, 23, 391–398.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244–249.

Brown, H. (1985). The effects of set content and temporal context on musician's aural perception of tonality. Unpublished doctoral dissertation, The Ohio State University. *Dissertation Abstracts International*, 46, 3527A.

Brown, H. (1987). Tonal hierarchies and perceptual context: An experimental study of musical behaviour. *Psychomusicology* , 7(1), 77–90.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, 5, 219–250.

Brown, H., & Butler, D. (1981). Diatonic trichords as minimal tonal cue–cells. *In Theory Only*, 5, 39–55.

Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only*, 5, 37–55.

Browne, R. (1985). The dialectic of good continuation in tonal music. *Music Analysis*, 4, 5–13.

Bruning, J. L., & Kintz, B. L. (1977). *Computational Handbook of Statistics*. Illinois: Scott, Foresman.

Bryman, A., & Cramer, D. 1990. *Quantitative data analysis for social scientists*. London: Routledge.

Burns, E. M., & Ward, W. D. (1982). Intervals, scales and tuning. In D. Deutsch (Ed.), The *Psychology of Music* (*pp.* 241–269). London: Academic Press.

Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, 6, 219–242.

Butler, D. (1990). A study of event hierarchies in tonal and post–tonal music. *Psychology of Music*, 18, 4–17.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music or North India. *Journal of Experimental Psychology: General*, 113, 394–412.

Clarke, E. (1989). Using computers in psychology of music teaching and research. *Musicus*, 1, 17–21.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, 8, 240–247.

Cook, N. (1987a). Review of *Music Perception* (Ed. Diana Deutsch), Vol. 1, Nos. 1–4; Vol. 2, Nos. 1–2. *Music Analysis*, 6, 168–178.

Cook, N. (1987b). The perception of large–scale tonal closure. *Music Perception*, 5, 197–206.

Croonen, W. L. M., & Kop, P. F. M. (1989). Tonality, tonal scheme, and contour in delayed recognition of tone sequences. *Music Perception*, 7, 49–68.

Cross, I. (1985). Music and change: On the establishment of rules. In P. Howell, I. Cross & R. West (Eds.), *Musical structure and cognition* (*pp*. 1–20). London: Academic Press.

Cross, I., Howell, P., & West, R. (1983). Preferences for scale structure in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 444–460.

Cross, I., Howell. P., & West, R. (1985). Structural relationships in the perception of musical pitch. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (*pp.* 121–142). London: Academic Press.

Cross, I., West, R., & Howell, P. (1985). Pitch relations and the formation of scalar structure. *Music Perception*, 2, 329– 344.

Cross, I., West, R., & Howell, P. (1991). Cognitive correlates of tonality. In P. Howell, R. West, & I. Cross (Eds.), *Representing Musical Structure*. London: Academic Press.

Cuddy, L. L. (1985). The color of melody. *Music Perception*, 2, 345–360.

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics*, 41, 609–620.

Deutsch, D. (1969). Music Recognition. *Psychological Review*, 76, 300–307.

Deutsch, D. (Ed). (1982). *The Psychology of Music*. New York: Academic Press.

Deutsch, D. & Feroe, J. (1981). The internal representation of pitch structure in tonal music. *Psychological Review*, 88, 503–522.

Dosher, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of*

Experimental Psychology: General, 118(2), 191–211.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85, 341–354.

Dowling, W. J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), *The Psychology of Music* (*pp*. 413–429). London: Academic Press.

Drobisch (1846) cited by Ruckmick, C. A. (1929). A new classification of tonal qualities. *Psychological Review*, 36, 172–180.

Edworthy, J. (1983). *An experimental investigation into the relationship between pitch–interval and contour in melody processes.* Unpublished doctoral dissertation, University of Warwick, England.

Edworthy, J. (1985a). Interval and contour in melody processing. *Music Perception*, 3, 375–388.

Edworthy, J. (1985b). Melodic contour and musical structure. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (*pp*. 169–188). London: Academic Press.

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, 1, 245–252.

Firebaugh, M. W. (1989). *Artificial Intelligence: A Knowledge–Based Approach.* Boston: PSW–Kent.

Fiske, H. E. (1982a). Chronometric analysis of selected pattern discrimination tasks in music listening. *Psychology of Music*, 10, 37–47.

Fiske, H. E. (1982b). The application of stage reduction theory to music listening. *Psychology of Music*, special issue, 31–35.

Fiske, H. E. (1985). Cognition strategies in music listening. *Bulletin of the Council for Research in Music Education*, 85, 56–64.

Fiske, H. E. (1987). Cognition structure and the perception of music. *Bulletin of the Council for Research in Music Education*, 91, 31–37.

Fiske, H. E. (1990). *Music and Mind.* Lampeter: Edwin Mellen.

Galton, F. (1899). On instruments for (1) testing perception of differences for tint and for (2) determining reaction time. *Journal of the Anthropological Institute*, 19, 27–29.

Guildford, J. P., & Fruchter, B. 1978. *Fundamental statistics in psychology and education.*. London: McGraw–Hill.

Hahn, J., & Jones, M. R. (1981). Invariants in auditory frequency relations. *Scandinavian Journal of Psychology*, 22, 129–144.

Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Trans.). New York, Dover. (Original work published 1863.)

Hodgkins, J. (1962). Influence of age on the speed of reaction and movement in females. *Journal of Gerontology*, 17, 385–389.

Howell, P., West, R., & Cross, I. (1984). The detection of notes incompatible with scalar structure. *Journal of the Acoustical Society of America*, 76, 1682–1689.

Hubbard, T. L., & Stoeckig, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(4), 656–667.

Hugin, F., Norris, A. H., & Shock, N. W. (1960). Skin reflex and voluntary reaction times in young and old males. *Journal of Gerontology*, 15, 388–391.

Imberty, M. (1969). *L'acquisition des structures tonales chez l'enfant*. Paris: Klincksieck.

Janata, P., & Reisberg, D. (1988). Response–time measures as a means of exploring tonal hierarchies. *Music Perception*, 6, 161–172.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32, 241–254.

Jones, M. R. (1981). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology* , 1, 34–51.

Jones, M. R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model. *Psychomusicology* , 2, 1–13.

Jordan, D. S. (1987). Influence of the diatonic tonal hierarchy at microtonal intervals. *Perception & Psychophysics*, 41, 482–488.

Jordan, D. S., & Shepard, R. N. (1987). Tonal schemas: Evidence obtained by probing distorted musical scales. *Perception & Psychophysics*, 41, 489–504.

Juola, J. F., Schadler, M., Chabot, R. J., & McCaughey, M. W. (1978). The development of visual information processing skills related to reading. *Journal of Experimental Child Psychology*, 25, 459–476.

Kail, R. (1988). Developmental functions for speeds of cognitive processes. *Journal of Experimental Child Psychology*, 45, 339–364.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception & Psychophysics*, 32, 37–49.

Kelley, Z. A., & Brandt, J. F. (1984). Pitch change recognition as a function of duration in successive dichotic stimuli. *Psychology of Music*, 12, 43–59.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15, 166–177.

Kohonen, T. (1984). *Self–organisation and Associative Memory*. Berlin: Springer–Verlag.

Kristofferson , M. W. (1972a). Effects of practice on character–classification performance. *Canadian Journal of Psychology*, 26, 540–560.

Kristofferson , M. W. (1972b). When item recognition and visual search functions are similar. *Perception & Psychophysics*, 12, 379–384.

Krueger, L. E. (1978). A theory of perceptual matching. *Psychological Review*, 85, 278–304.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 346–374.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception*, 1, 28–62.

Krumhansl, C. L. (1990a). *Cognitive foundations of musical pitch.* Oxford: Oxford University Press.

Krumhansl, C. L. (1990b). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, 7, 309-324.

Krumhansl, C. L., Bharucha, J. J., & Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics*, 32, 96–108.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 24–36.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in Music Perception. *Memory & Cognition*, 11, 325–334.

Krumhansl, C. L., & Keil, F. C., (1982). Acquisition of the hierarchy of tonal functions in music, *Memory and Cognition*, 10, 243–251.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, 89, 334–368.

Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve–tone serial music. *Music Perception*, 5, 31–78.

Krumhansl, C. L., & Schmuckler, M. A. (1986). The Petrushka Chord: A perceptual investigation. *Music Perception*, 4, 153–184.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579–594.

Kruskal, J. B. (1964a). Multidimensional scaling by optimising goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1–27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29, 28–42.

Leman, M. (1991a). Artificial neural networks in music research. In A. Marsden & A. Pople (Eds.), *Computer representations and models in music*. London: Academic Press.

Leman, M. (1991b). The ontogenesis of tonal semantics: results of a computer study. In P. M. Todd & D. G. Loy (Eds.), *Music and Connectionism* (pp. 100-127). Cambridge, Massachusetts, MIT Press.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception*, 5, 315–350.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.

Meyer, L. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review*, 95(2), 183–237.

Morrongiello, B. A. (1992). Effects of training on children's perception of music: A review. *Psychology of Music*, 20, 29–41.

Morrongiello, B. A., & Roes, C. L. (1990). Developmental changes in children's perception of musical sequences: Effects of musical training. *Developmental Psychology*, 26(5), 814–820.

Nettelbeck, T, & Brewer, N. (1976). Effects of stimulus–response variables on the choice reaction time of mildly retarded adults. *American Journal of Mental Deficiency*, 81, 85–92.

Nickerson, R. S. (1972). Binary–classification reaction time: A review of some studies of human information processing capabilities. *Psychonomic Monograph Supplements*, 4(17), 275-318.

Noble, C.E., Baker, B. L., & Jones, T. A. (1964). Age and sex parameters in psychomotor learning. *Perceptual and Motor Skills*, 19, 935–945.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition (*pp*. 41–82). New York: Halstead Press.

Palmer, C., & Krumhansl, C. L. (1987a). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics*, 41, 505–518.

Palmer, C., & Krumhansl, C. L. (1987b). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 116– 126.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441–474.

Pechstedt, P. H., Kershner, J., & Kinsbourne, M. (1989). Musical training improves processing of tonality in the left hemisphere. *Music Perception*, 6, 275–298.

Proctor, R. W., Healy, A. F., & Zandt, T. V. (1991). Same–different judgments of multiletter strings: insensitivity to positional bias and spacing. *Perception & Psychophysics*, 49(1), 62–72.

Proctor, R. W., & Rao, K. V. (1982). On the 'misguided' use of reaction–time differences: A discussion of Ratcliff and Hacker (1981). *Perception & Psychophysics*, 31, 601–602.

Proctor, R. W., Zandt, T. V., & Watson, H. D. (1990). Effects of background symmetry on same–different pattern matching: A compromise– criteria account. *Perception & Psychophysics*, 48(6), 543–550.

Ratcliff, R. (1978). A theory of order mental retrieval. *Psychological Review*, 85, 59–108.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, 88, 552–572.

Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, 95, 238–254.

Ratcliff, R., & Hacker, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception &*

Psychophysics, 30, 303–307.

Ratcliff, R., & Hacker, M. J. (1982). On the misguided use of reaction–time differences: A reply to Procter and Rao (1982). *Perception & Psychophysics*, 31, 603–604.

Revesz, G. (1954). *An introduction to the Psychology of Music*. Norman, Oklahoma: Oklahoma University Press.

Roberts, L. A. (1986). Modality and suffix effects in memory for melodic and harmonic musical materials. *Cognitive Psychology*, 18, 123–157.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532–547.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.

Ruckmick, C. A. (1929). A new classification of tonal qualities. *Psychological Review*, 36, 172–180.

Scarborough, D. L., Miller, B. O., & Jones, J. A. (1989). Connectionist models for tonal analysis. *Computer Music Journal*, 13, 49–55.

Scarborough, D. L., Miller, B. O., & Jones, J. A. (1991). Connectionist models for tonal analysis. Addendum in P. M. Todd & D. G. Loy (Eds.), *Music and Connectionism* (*pp.* 61-63). Cambridge, Massachusetts, MIT Press.

Schenker, H. (1954). *Harmony* (O. Jones, Ed. and E. M. Borgese, Trans.). Cambridge. Massachusetts: MIT Press. (Originally published, 1906).

Schenker, H. (1979). *Free composition* (E. Oster, Ed. and Trans.). New York: Longman. (Originally published , 1935).

Schoenberg, A. (1969). *Structural functions of harmony* (revised edition). New York: Norton.

Schoenberg, A. (1978). *Theory of harmony* (R. Carter, Trans.). London: Faber & Faber. (originally published, 1911.)

Schouten, J. F. (1938). The perception of subjective tones. *Proceedings of the Koninklijke Nederlands Akademie van Wetenschappen*, 41, 1083–1093.

Seebeck, A. (1841). Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Annalen der Physik und Chemie*, 53, 417–436.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. & II. *Psychometrika*, 27, 125–140, 219–246.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2346–2353.

Shepard, R. N. (1982a). Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89, 305–333.

Shepard, R. N. (1982b). Structural representations of musical pitch. In D. Deutsch (Ed.), The *Psychology of Music* (*pp*. 343–390). London: Academic Press.

Simon, H. A. (1968). Perception du Pattern Musical par AUDITEUR.

Science de l'art. Tome, 2, 28–34.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 457–471.

Speer, J. R., & Adams, W. E. (1985). *The effects of musical training upon the development of the perception of musical pitch.* Paper presented at the meetings of the Society for Research in Child Development, Toronto, April 1985.

Speer, J. R., & Meeks, P. U. (1985). School children's perception of pitch in music. *Psychomusicology* , 5, 49–56.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276–315). Amsterdam: North–Holland.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27, 195–232.

Stevens, S. S., Volkman, J. (1940).The relation of pitch to frequency: a revised scale. *American Journal of Psychology*, 53, 329–353.

Stevens, S. S., Volkman, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America*, 8, 185–190.

Suffolk Reading Scale. (1981). NFER: Nelson.

Swanwick, K. (1973). Musical cognition and aesthetic response. *Psychology of Music*, 1, 7–13.

Temko, P. M. (1971). *The perception of pitch predominance in selected musical examples of avant-garde composers: 1945–1961.* Unpublished doctoral dissertation, Florida State University. Dissertation Abstracts International, 32, 5275A–5276A.

Teplov, B. M. (1966). *Psychologie des aptitudes musicales.* Paris: Presses Universitaires de France.

Terhardt, E. (1976). Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz. *Acustica*, 36, 121–137.

Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, 23, 483–492.

Trehub, S. E. (1987). Infants' perception of musical patterns. Perception and Psychophysics, 41, 635–641.

Trehub, S. E., Cohen, A. J., Thorpe, L. A., & Morrongiello, B. A. (1986). Development of the perception of musical relations: semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 295–301.

Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences.* Unpublished doctoral dissertation, Eindhoven, Holland: Institute for Perception Research.

Van Noorden, L. P. A. S. (1982). Two channel pitch–perception. In M. Clynes (Ed.), *Music, mind and brain: the neuropsychology of music* (*pp.* 251–269). Plenum: New York.

Ward, J. D. (1984). *Music and imagination: children's constructs.* Unpublished doctoral dissertation, London University.

Watanabe, T. (1990). Effect of irrelevant differences as a function of the relations between relevant and irrelevant dimensions in the same–different task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(1), 132–142.

Welch, G. F. (1979). Vocal range and poor pitch singing. *Psychology of Music*, 7, 13–31.

Welch, G. F., Howard, D. M., & Rush, C. (1989). Real–time visual feedback in the development of vocal pitch accuracy in singing. *Psychology of Music*, 17, 146–157.

Welford, A. T. (1980). Relationships between reaction time and fatigue, stress age and sex. In A. T. Welford (Ed.), *Reaction Times (pp.* 321–354). London, Academic Press.

Wertheimer, M. (1955). Laws of organisation in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology*. London: Routledge & Kegan Paul. (First published in German, 1923.)

Williams, A.& Weisstein, N. (1978). Line segments are better perceived better in a coherent context than alone: An object–line effect in visual perception. *Memory and Cognition*, 6, 85–90.

Zenatti, A. (1969). Le développement génétique de la perception musicale. Monographies Francaises de Psychologie, 17, 39–48.

APPENDIX I

Technical Discussion of Computer Systems

The advantages and disadvantages of the BBC microcomputer, Atari ST and Hybrid Music System were investigated.

1. Atari ST computer

Although the BBC microcomputer was chosen for the experiments, the Atari ST microcomputer had been investigated as a possible computer which could have been used. The Atari ST is a later generation of microcomputer which offers two advantages over the BBC microcomputer:

a) The Atari ST microcomputer has a midi interface built into the computer and a midi provision within the operating system. This means that there are direct commands available from within BASIC (an acronym for *Beginners All-Purpose Symbolic Instruction Code*) which can make MIDI programming relatively easy. Unfortunately, the BBC microcomputer has no provision within the operating system for MIDI interaction. However, commands are available from BBC BASIC to output information to an attached MIDI interface and since one-way midi transmission was envisaged in the experiments (i.e. computer to sound module) either computer was equally suitable for the task.

b) The timing resolution available from BASIC on Atari computers is recorded in two-hundredths of a second by a system variable and may be found by peeking an address in the operating system workspace. This resolution is greater than the timing resolution of centi–seconds from BASIC on the BBC microcomputer. However, centi–seconds seemed a resolution measure of sufficient magnitude to discover differences of reaction time. The millisecond timing often used in studies of reaction time could possibly be implemented on microcomputer systems with specialist machine code servers which disable machine operating system interrupts. This was not possible in an experiment of this design as input–output operations via the operating system are not possible with interrupts disabled and consequently scanning of interfaced switches and driving of MIDI musical instruments is suspended. This would have made impossible the recognition of the pressing of switches on the button box and turning off a note on the keyboard if the switch had not been pressed. Millisecond timing was not readily available on either microcomputer. Both the BBC and Atari ST microcomputers have the provision of two joysticks which enable simple connection of a two-way button box so either computer was suitable for button-box connection.

2. BBC microcomputer

The BBC microcomputer, however, has three distinct advantages over the Atari ST for an experiment of this nature:

a) Mode 7 in BASIC (i.e. teletext mode) allows the mixing of double-height text and a palette of seven colours, whereas the standard (and available for the experiment) Atari monitor is monochrome (i.e. black and white). A colour environment was considered better for the children than a monochrome environment.

b) Programming in BASIC on the BBC is a better structured language than the poor BASIC supplied with the Atari. While it would have been possible to utilise disk-based *Fast Basic* or *'C'* on the Atari, these languages are not as available or as utilised as the ROM-based BBC BASIC. Furthermore, as *'C'* is compiled, development time would have been greatly extended as modifications would require complete re–compilation.

c) The BBC microcomputer is readily available in schools all over the country, particularly within the age-range of pupils who participated as subjects in the experiment. Utilisation of existing equipment in schools avoided transportation problems and shortened setting-up time.

3. Hybrid Music System

The Hybrid Music system as an extension of the BBC microcomputer had been contemplated as a possible programming environment for the experiment, although the AMPLE (Advanced Music Programming Language Environment) language offered no advantage in either in centi–second resolution for measuring reaction time responses or in the possibility of interfacing to midi instruments (via the Music 2000). The sound module (Music 5000) which is used by the Hybrid Music System is similar in sound quality to many midi sound modules. The system therefore offered no improvement in sound quality or improvement in data collection.

The pitch resolution of the Hybrid System is superior to other systems in providing greater control of intervals of less than a semitone than midi systems. However, this was unnecessary in the experiments reported here and consequently offered no advantage.

The Hybrid System had been used in the preparation of some experimental materials concerned with testing children's discrimination of less than a semitone in short musical excepts from various musical styles (*e.g.* Bartok, Hindemith), although the tests had not been further developed since the problem of musical context within such a diverse array of musical styles had proved problematic and led to the investigation of tonality as a meaningful cognitive construct for children.

4. Details of Equipment

The experimental situation comprised a standard BBC computer system with monitor and disk drive, button box, K1 MIDI interface and Yamaha PSS 480 MIDI keyboard.

The button box was manufactured with two switches linked to the two joystick ports on the D–connector on the BBC.

The K1 MIDI interface was obtained from: ESP, Holly Tree Cottage, Main Street, Strelley Village, Nottingham, NG8 6PD.

APPENDIX II

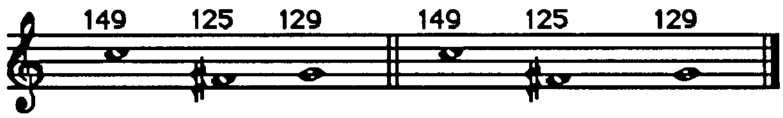
Pilot Study Materials (Chapter 4)

Music Notation

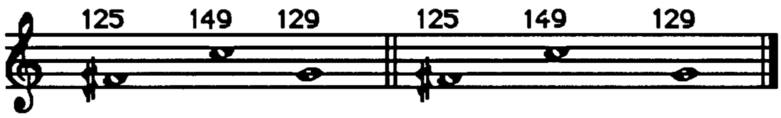
The following is a musical representation of the pitches presented for each of the twelve trials by the computer program in Appendix III.

Numbers above the notes relate to the numbers in the data statements in the computer program. These numbers are required by the BBC BASIC Sound command to represent these pitches

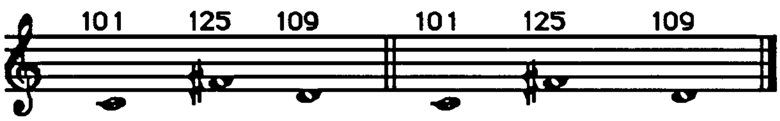
Trial 1



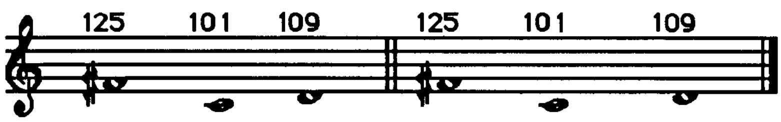
Trial 2



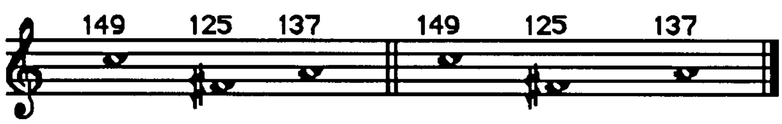
Trial 3



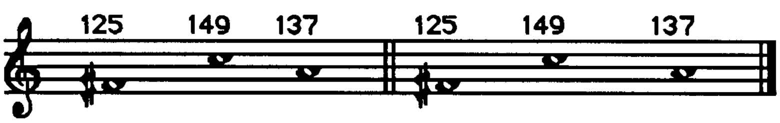
Trial 4



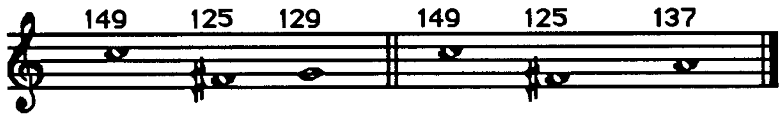
Trial 5



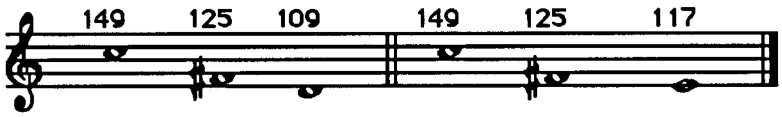
Trial 6



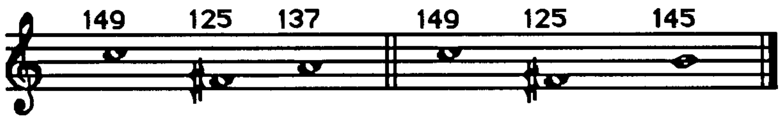
Trial 7



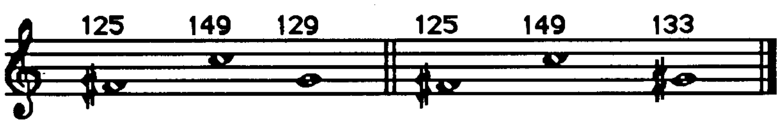
Trial 8



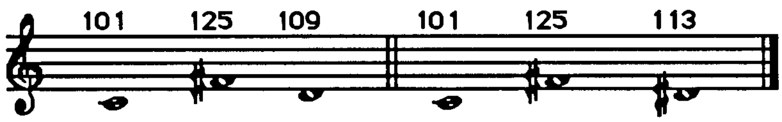
Trial 9



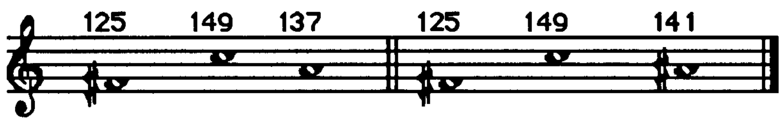
Trial 10



Trial 11



Trial 12



APPENDIX III

Pilot Study Materials (Chapter 4)

BBC Microcomputer Program

The following is a BASIC listing of the computer program which was written to present the experimental materials and record reaction time responses.

The musical pitches which it presents are notated in Appendix II.

The program may be found on the accompanying diskette as "$.PILOT".

```
10 *KEY9SAVE"PILOT"|M
20 :
30 :
40 :
50 REM ***** TEST 2.2.89 *****
60 :
70 REM ***** BBC Microcomputer *****
80 :
90 REM ***** by Richard E. Hodges *****
100 :
110 REM ***** INITIALISE *****
120 :
130 MODE 7
140 Q=12:REM ***** NUMBER OF QUESTIONS *****
150 DIM A(12):DIM B(12):DIM C(12):DIM D(12):DIM N(12)
160 FOR I=1 TO 12
170 READ A(I):READ B(I):READ C(I):READ D(I)
180 NEXT I
190 :
200 :
210 REM ***** GENERATE RANDOM NUMBERS *****
220 FOR G=1 TO 12
230 R=INT(RND(1)*12+1)
240 F%=0
250 FOR T=1 TO G
260 IF N(T)=R THEN F%=1
270 NEXT T
280 N(G)=R
290 IF F%=1 THEN GOTO 230
300 NEXT G
310 :
```

```
320 REM ***** OPEN FILE FOR RESPONSES *****
330 :
340 CLS
350 PRINT TAB(3,7);CHR$(130);"WHAT IS YOUR NAME   ";:INPUT A$
360 KF=OPENOUT(A$)
370 :
380 REM ***** INSTRUCTIONS *****
390 :
400 T=TIME:PAUSE=500
410 CLS
420 PRINT TAB(3,5);CHR$(131);"YOU WILL HEAR TWO 3-NOTE TUNES"
430 REPEAT UNTIL TIME=T+PAUSE
440 PRINT TAB(7,9);CHR$(131);"THEY MAY BE THE ";CHR$(130);"SAME"
450 REPEAT UNTIL TIME=T+(PAUSE*2)
460 PRINT TAB(5,11);CHR$(131);"OR THEY MAY BE
";CHR$(130);"DIFFERENT"
470 REPEAT UNTIL TIME=T+(PAUSE*3)
480 CLS
490 PRINT TAB(2,5)"PRESS  ";CHR$(130);"'S'";CHR$(135);" IF THEY ARE THE
SAME"
500 REPEAT UNTIL TIME=T+(PAUSE*4)
510 PRINT TAB(2,9)"PRESS  ";CHR$(130);"'D'";CHR$(135);" IF THEY ARE
DIFFERENT"
520 REPEAT UNTIL TIME=T+(PAUSE*5)
530 CLS:PRINT TAB(5,20) "PRESS SPACE-BAR TO BEGIN"
540 G=GET:REPEAT UNTIL (G=32)
550 CLS:T=TIME:REPEAT UNTIL TIME>T+200
560 :
570 REM ***** TEST LOOP *****
580 :
590 FOR L=1 TO Q
600 I=N(L)
610 CLS
620 T=TIME
630 SOUND 1,-15,A(I),8
640 SOUND 1,0,50,4
650 SOUND 1,-15,B(I),8
660 SOUND 1,0,50,4
670 SOUND 1,-15,C(I),8
680 SOUND 1,0,50,4
690 REPEAT UNTIL TIME>T+400
700 SOUND 1,-15,A(I),8
710 SOUND 1,0,50,4
720 SOUND 1,-15,B(I),8
730 SOUND 1,0,50,4
740 SOUND 1,-15,D(I),8
750 SOUND 1,0,50,4
760 :
770 REM ***** GET RESPONSE *****
780 :
790 REM S=68;D=83
800 :
810 T=TIME
820 REPEAT:G=GET:UNTIL (G=68) OR (G=83)
830 DELAY=TIME-T
840 IF G=68 THEN PRINT TAB(12,9);CHR$(129);CHR$(141);" DIFFERENT"
850 IF G=68 THEN PRINT TAB(12,10);CHR$(129);CHR$(141);" DIFFERENT"
860 IF G=83 THEN PRINT TAB(12,9);CHR$(131);CHR$(141);" SAME"
870 IF G=83 THEN PRINT TAB(12,10);CHR$(131);CHR$(141);" SAME"
```

```
880 PRINT#KF,I
890 PRINT#KF,A(I)
900 PRINT#KF,B(I)
910 PRINT#KF,C(I)
920 PRINT#KF,D(I)
930 PRINT#KF,DELAY
940 IF G=68 THEN PRINT#KF,"DIFFERENT"
950 IF G=83 THEN PRINT#KF,"SAME"
960 :
970 REM ***** DATA STATEMENTS *****
980 :
990 REM ***** SAME STIMULI *****
1000 :
1010 DATA 149,125,129,129
1020 DATA 125,149,129,129
1030 DATA 101,125,109,109
1040 DATA 125,101,109,109
1050 DATA 149,125,137,137
1060 DATA 125,149,137,137
1070 :
1080 REM ***** DIFFERENT STIMULI *****
1090 :
1100 DATA 149,125,129,137
1110 DATA 149,125,109,117
1120 DATA 149,125,137,145
1130 DATA 125,149,129,133
1140 DATA 101,125,109,113
1150 DATA 125,149,137,141
1160 :
1170 :
1180 REM ***** PRESS SPACE FOR NEXT ITEM *****
1190 :
1200 T=TIME:REPEAT UNTIL TIME>T+200
1210 PRINT TAB(2,20)."PRESS SPACE BAR FOR NEXT TWO TUNES"
1220 REPEAT:G=GET:UNTIL (G=32)
1230 CLS:T=TIME:REPEAT UNTIL TIME>T+200
1240 NEXT L
1250 CLS
1260 PRINT TAB(5,9);CHR$(131);CHR$(141);"THANK YOU FOR YOUR HELP"
1270 PRINT TAB(5,10);CHR$(131);CHR$(141);"THANK YOU FOR YOUR HELP"
1280 FOR M=1 TO 2000:NEXT
1290 CLOSE#KF
1300 END
1310 :
1320 :
1330 REM ***** TEST RANDOM NUMBERS *****
1340 FOR I=1 TO 12:PRINT N(I):NEXT
```

APPENDIX IV

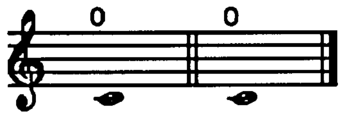
Experiment One Materials (Chapter 5)

Music Notation

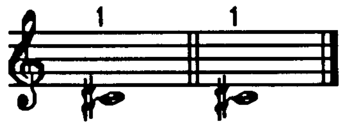
The following is a musical representation of the pitches presented for each of the twenty trials by the computer program in Appendix V.

Numbers above the notes relate to the numbers in the data statements in the computer program. These numbers represent semitone offsets from the base note which is middle C (e.g. C=0, D flat=1, D=2, etc.).

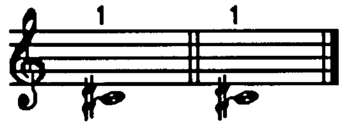
Trial 1



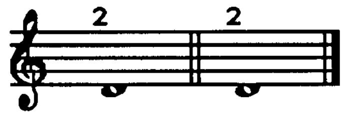
Trial 2



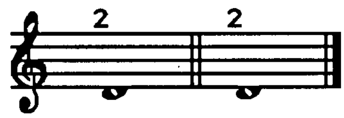
Trial 3



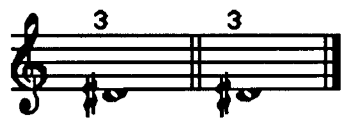
Trial 4



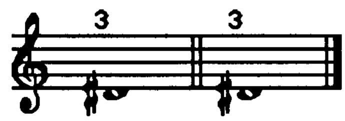
Trial 5



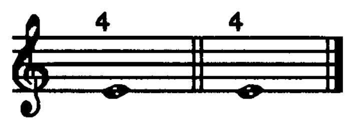
Trial 6



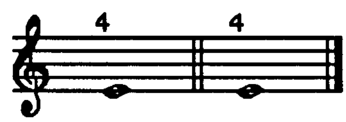
Trial 7



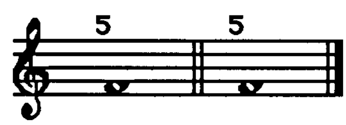
Trial 8



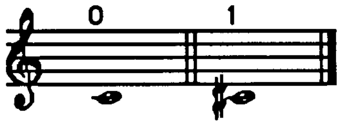
Trial 9



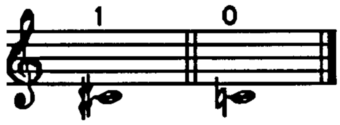
Trial 10



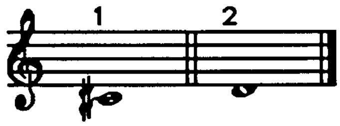
Trial 11



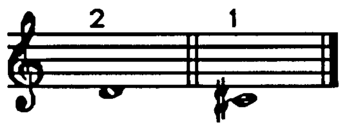
Trial 12



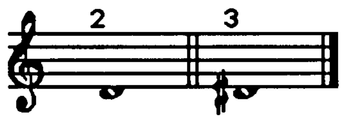
Trial 13



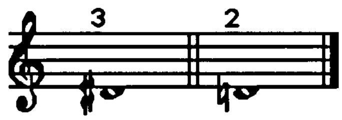
Trial 14



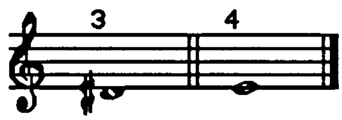
Trial 15



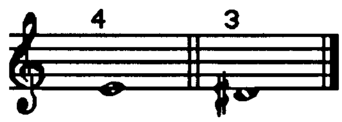
Trial 16



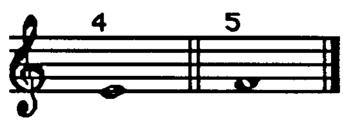
Trial 17



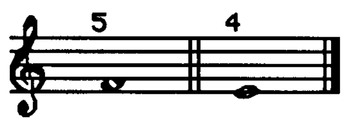
Trial 18



Trial 19



Trial 20



APPENDIX V

Experiment One Materials (Chapter 5)

BBC Microcomputer Program

The following is a BASIC listing of the computer program which was written to present the experimental materials and record reaction time responses.

The musical pitches which it presents are notated in Appendix IV.

The program may be found on the accompanying diskette as "$.MB1".

```
10 *KEY0RUN|M
20 *KEY1LIST|M
30 *KEY9SAVE"MB1"|M
40 :
50 ON ERROR GOTO 3010
60 :
70 REM ***** REACTION TIME TEST version 28.1.91 *****
80 REM ***** BBC Microcomputer *****
90 REM ***** by Richard E. Hodges *****
100 :
110 REM ***** USER INPUT *****
120 REM ***** DATE *****
130 DATE$="31.1.91"
140 REM ***** TEST DESCRIPTOR *****
150 TEST$="TEST ONE"
160 REM ***** NUMBER OF QUESTIONS *****
170 Q=20
180 REM ***** PAUSE TIME FOR INSTRUCTIONS *****
190 REM ***** SHOULD BE 199 *****
200 PAUSE=199
210 :
220 REM ***** PROCEDURE CALLS *****
230 MODE 7
240 PROCinitialise
250 PROCtitle
260 PROCrandom
270 PROCgetname
280 PROCinstructions
```

```
290 PROCtestloop
300 PROCthankyou
310 PROCwritedata
320 PROCtitle
330 PROCtestend
340 PROCreset
350 END
360 :
370 :
380 REM ***** TITLE PAGE *****
390 DEFPROCtitle
400 CLS
410 REM ***** TURN OFF CURSOR *****
420  VDU 23,1,0,0,0,0,0,0,0
430 FOR I=1 TO 106
440 PROCrandomcolour
450 PRINT CHR$(COL);"MUSIC TEST   ";
460 NTE=40+I/2
470 PROCnoteon
480 FOR D=1 TO RND(30)+10
490 NEXT D
500 PROCnoteoff
510 NEXT I
520 ENDPROC
530 :
540 REM ***** INITIALISE *****
550 DEFPROCinitialise
560 REM ***** SET MIDI INTERFACE *****
570 ?&FC00=&3:?&FC00=&15
580 :
590 REM ***** DISABLE AUTO REPEAT DELAY ON KEYBOARD *****
600 *FX 11,0
610 :
620 REM ***** A TO D CONVERTER OFF *****
630 *FX 16,0
640 :
650 REM ***** TURN OFF CURSOR *****
660  VDU 23,1,0,0,0,0,0,0,0
670 :
680 REM ***** SET COLOURS *****
690 RED$=CHR$(129)
700 GREEN$=CHR$(130)
710 YELLOW$=CHR$(131)
720 BLUE$=CHR$(132)
730 MAGENTA$=CHR$(133)
740 CYAN$=CHR$(134)
750 WHITE$=CHR$(135)
760 DH$=CHR$(141)
770 :
780 REM ***** DIMENSION ARRAYS *****
790 DIM A(Q):DIM B(Q):DIM N(Q)
800 DIM DELAY(Q): DIM RESPONSE$(Q)
810 FOR I=1 TO Q
820 READ A(I):READ B(I)
830 NEXT I
840 ENDPROC
850 :
860 REM ***** RANDOM COLOUR *****
870 DEFPROCrandomcolour
```

```
880 COL=128+RND(7)
890 ENDPROC
900 :
910 REM ***** GENERATE RANDOM NUMBERS *****
920 DEFPROCrandom
930 FOR G=1 TO Q
940 R=INT(RND(1)*Q+1)
950 F%=0
960 FOR T=1 TO G
970 IF N(T)=R THEN F%=1
980 NEXT T
990 N(G)=R
1000 IF F%=1 THEN GOTO 940
1010 NEXT G
1020 ENDPROC
1030 :
1040 REM ***** GET NAME FOR RESPONSES *****
1050 DEFPROCgetname
1060 CLS
1070 DOWN=7:TEXT$=CYAN$+"WHAT IS YOUR NAME ?"
1080 PROCcentretext
1090 A=0:NAME$=""
1100 REPEAT
1110 G=GET:A=A+1
1120 NAME$=NAME$+CHR$(G)
1130 DOWN=11:TEXT$=DH$+GREEN$+NAME$+"  "
1140 PROCcentretext
1150 UNTIL G=13 OR G=127
1160 IF A<2 OR G=127 THEN 1060
1170 CLS
1180 DOWN=14:TEXT$=" Pleased to meet you, "+NAME$
1190 PROCcentretext
1200 PROCpause
1210 PROCpressspace
1220 ENDPROC
1230 :
1240 REM ***** INSTRUCTIONS *****
1250 DEFPROCinstructions
1260 CLS
1270 DOWN=9:TEXT$=WHITE$+"YOU WILL
HEAR"+GREEN$+"TWO"+WHITE$+"NOTES"
1280 PROCcentretext
1290 PROCpause
1300 PROCpressspace
1310 DOWN=9:TEXT$=RED$+"THEY MAY BE THE "+YELLOW$+"SAME"
1320 PROCcentretext
1330 PROCpause
1340 PROCpressspace
1350 DOWN=9:TEXT$=RED$+"OR THEY MAY BE "+BLUE$+"DIFFERENT"
1360 PROCcentretext
1370 PROCpause
1380 PROCpressspace
1390 DOWN=5:TEXT$="PRESS "+YELLOW$+"'S'"+WHITE$+" IF THEY ARE
THE SAME"
1400 PROCcentretext
1410 PROCpause:PROCpause
1420 DOWN=9:TEXT$="PRESS "+BLUE$+"'D'"+WHITE$+" IF THEY ARE
DIFFERENT"
1430 PROCcentretext
```

```
1440 PROCpause
1450 DOWN=20:TEXT$=MAGENTA$+"PRESS SPACE-BAR TO BEGIN"
1460 PROCcentretext
1470 REM ***** FLUSH KEYBOARD BUFFER *****
1480 *FX 21,0
1490 G=0:REPEAT:G=GET:UNTIL G=32
1500 CLS:PROCpause
1510 ENDPROC
1520 :
1530 REM ***** CENTRE TEXT *****
1540 DEFPROCcentretext
1550 IF LEN(TEXT$)/2<>LEN(TEXT$+" ")/2 THEN TEXT$=TEXT$+" "
1560 PRINT TAB(19-(LEN(TEXT$)/2),DOWN),DH$+TEXT$
1570 PRINT TAB(19-(LEN(TEXT$)/2),DOWN+1),DH$+TEXT$
1580 ENDPROC
1590 :
1600 REM ***** PAUSE *****
1610 DEFPROCpause
1620 T=TIME:REPEAT UNTIL TIME>T+PAUSE
1630 ENDPROC
1640 :
1650 REM ***** TEST LOOP *****
1660 DEFPROCtestloop
1670 FOR L=1 TO Q
1680 I=N(L)
1690 REM ***** 60 = Midi Note Number for Middle C *****
1700 BASE=60
1710 CLS
1720 PROCrandomcolour
1730 DOWN=3:TEXT$=CHR$(COL)+"Test Item Number "+STR$(L)
1740 PROCcentretext
1750 PROCpause
1760 T2=TIME
1770 NTE=A(I)+BASE
1780 PROCnoteon
1790 PROCdelay
1800 PROCnoteoff
1810 REPEAT UNTIL TIME>T2+299
1820 :
1830 NTE=B(I)+BASE
1840 PROCnoteon
1850 PROCdelay1
1860 REPEAT UNTIL TIME>T1+99
1870 PROCnoteoff
1880 IF B<1 OR B>2 THEN PROCtestbutton:PROCprintresponse
1890 PROCpressspace
1900 NEXT L
1910 ENDPROC
1920 :
1930 REM ***** TEST BUTTON *****
1940 DEFPROCtestbutton
1950 REPEAT
1960 B=ADVAL(0) AND 3
1970 UNTIL B=1 OR B=2
1980 DELAY(I)=TIME-T1
1990 ENDPROC
2000 :
2010 REM ***** PRINT REPONSE *****
2020 DEFPROCprintresponse
```

```
2030 TEXT$="Delay is "+STR$(DELAY(I))+" centi-seconds"
2040 REM PRINT TAB(20-LEN(TEXT$)/2,18);TEXT$
2050 IF B=1 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+BLUE$+"DIFFERENT":PROCcentretext
2060 IF B=2 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+YELLOW$+"SAME":PROCcentretext
2070 IF B=1 THEN RESPONSE$(I)="DIFFERENT"
2080 IF B=2 THEN RESPONSE$(I)="SAME"
2090 ENDPROC
2100 :
2110 REM ***** DATA STATEMENTS *****
2120 REM ***** SAME STIMULI *****
2130 DATA 0, 0
2140 DATA 1, 1
2150 DATA 1, 1
2160 DATA 2, 2
2170 DATA 2, 2
2180 DATA 3, 3
2190 DATA 3, 3
2200 DATA 4, 4
2210 DATA 4, 4
2220 DATA 5, 5
2230 REM ***** DIFFERENT STIMULI *****
2240 DATA 0, 1
2250 DATA 1, 0
2260 DATA 1, 2
2270 DATA 2, 1
2280 DATA 2, 3
2290 DATA 3, 2
2300 DATA 3, 4
2310 DATA 4, 3
2320 DATA 4, 5
2330 DATA 5, 4
2340 :
2350 REM ***** PRESS SPACE FOR NEXT ITEM *****
2360 DEFPROCpressspace
2370 T=TIME:REPEAT UNTIL TIME>T+99
2380 DOWN=20:TEXT$=CYAN$+"PRESS SPACE BAR TO CONTINUE"
2390 PROCcentretext
2400 REM ***** FLUSH KEYBOARD BUFFER *****
2410 *FX 21,0
2420 G=0:REPEAT:G=GET:UNTIL G=32
2430 CLS:T=TIME:REPEAT UNTIL TIME>T+199
2440 ENDPROC
2450 :
2460 REM ***** MIDI NOTE ON *****
2470 DEFPROCnoteon
2480 ?&FC01=&90:?&FC01=NTE:?&FC01=64
2490 ENDPROC
2500 :
2510 REM ***** MIDI NOTE OFF *****
2520 DEFPROCnoteoff
2530 ?&FC01=&80:?&FC01=NTE:?&FC01=0
2540 ENDPROC
2550 :
2560 REM ***** DELAY *****
2570 DEFPROCdelay
2580 T=TIME:REPEAT UNTIL TIME>T+99
2590 ENDPROC
```

```
2600 :
2610 REM ***** DELAY1 *****
2620 DEFPROCdelay1
2630 T1=TIME
2640 REPEAT
2650 B=ADVAL(0) AND 3
2660 UNTIL B=1 OR B=2 OR TIME>T1+99
2670 IF B=1 OR B=2 THEN DELAY(I)=TIME-T1:PROCprintresponse
2680 IF B=3 THEN DOWN=7:TEXT$="DON'T PRESS BOTH BUTTONS
TOGETHER":PROCcentretext
2690 ENDPROC
2700 :
2710 REM ***** THANK YOU *****
2720 DEFPROCthankyou
2730 CLS
2740 DOWN=9:TEXT$=MAGENTA$+"THANK YOU FOR YOUR HELP"
2750 PROCcentretext
2760 ENDPROC
2770 :
2780 REM ***** WRITE DATA TO FILE *****
2790 DEFPROCwritedata
2800 RESULTS=OPENOUT("R."+NAME$)
2810 PRINT #RESULTS,DATE$,TEST$,Q
2820 FOR I=1 TO Q
2830 PRINT #RESULTS,N(I)
2840 NEXT I
2850 FOR I=1 TO Q
2860 PRINT #RESULTS,A(I),B(I)
2870 PRINT #RESULTS,DELAY(I),RESPONSE$(I)
2880 NEXT I
2890 CLOSE #RESULTS
2900 PROCpause
2910 ENDPROC
2920 :
2930 REM ***** TEST ENDING *****
2940 DEFPROCtestend
2950 REM ***** FLUSH KEYBOARD BUFFER *****
2960 *FX 21,0
2970 G=0:REPEAT:G=GET:UNTIL G<>0
2980 CLS
2990 ENDPROC
3000 :
3010 REM ***** ERROR HANDLING ROUTINE *****
3020 CLS: REPORT
3030 PRINT " AT LINE ";ERL
3040 PROCreset
3050 END
3060 :
3070 REM ***** RESET KEYBOARD AUTO-REPEAT RATE *****
3080 DEFPROCreset
3090 *FX 12,0
3100 ENDPROC
3110 :
3120 REM ***** TEST RANDOM NUMBERS *****
3130 FOR I=1 TO Q
3140 PRINT N(I)
3150 NEXT I
3160 END
3170 :
```

```
3180 REM ***** TEST NOTES IN ARRAY *****
3190 FOR I=1 TO Q
3200 PRINT A(I),B(I)
3210 NEXT I
3220 END
```

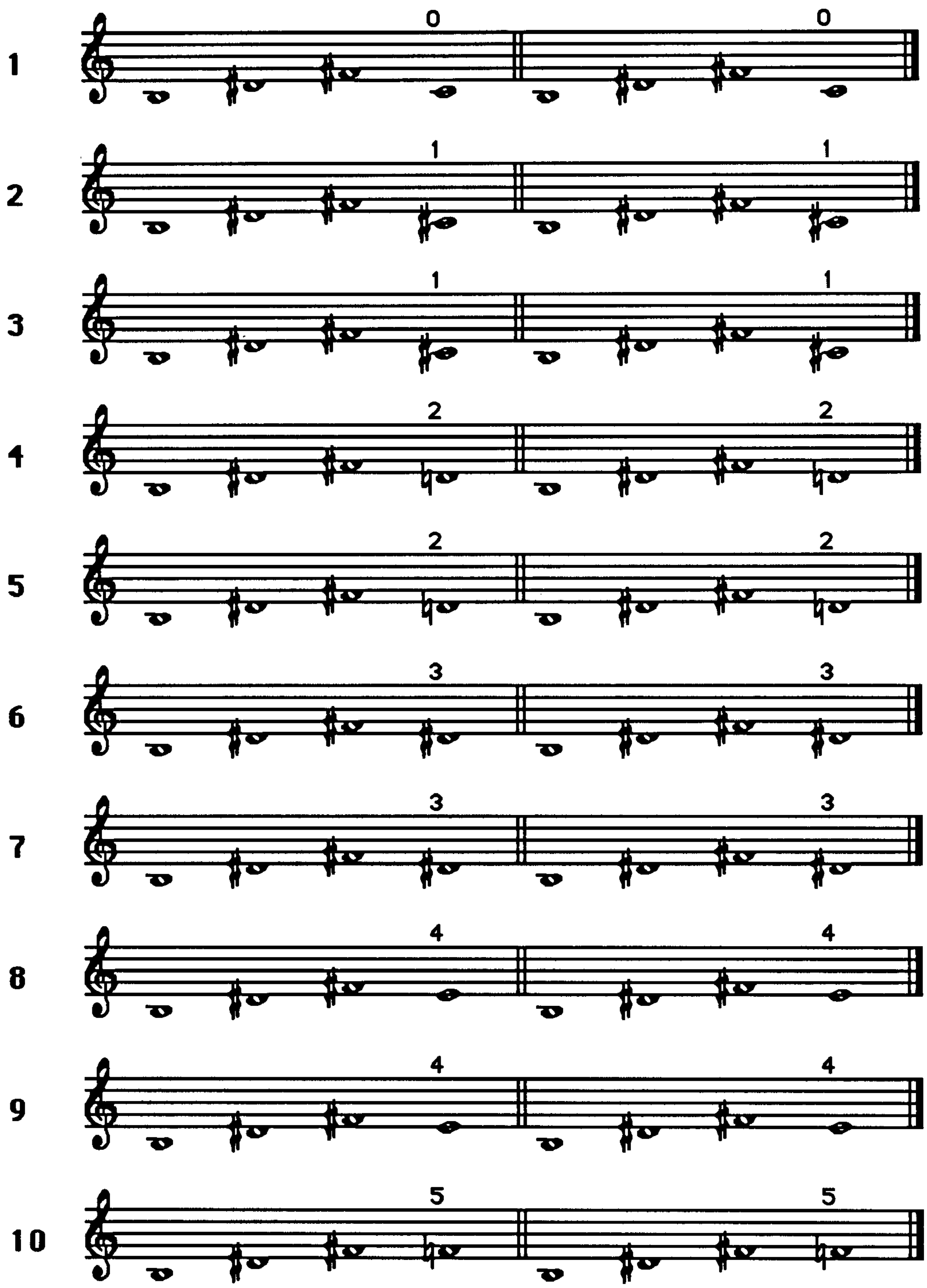
APPENDIX VI

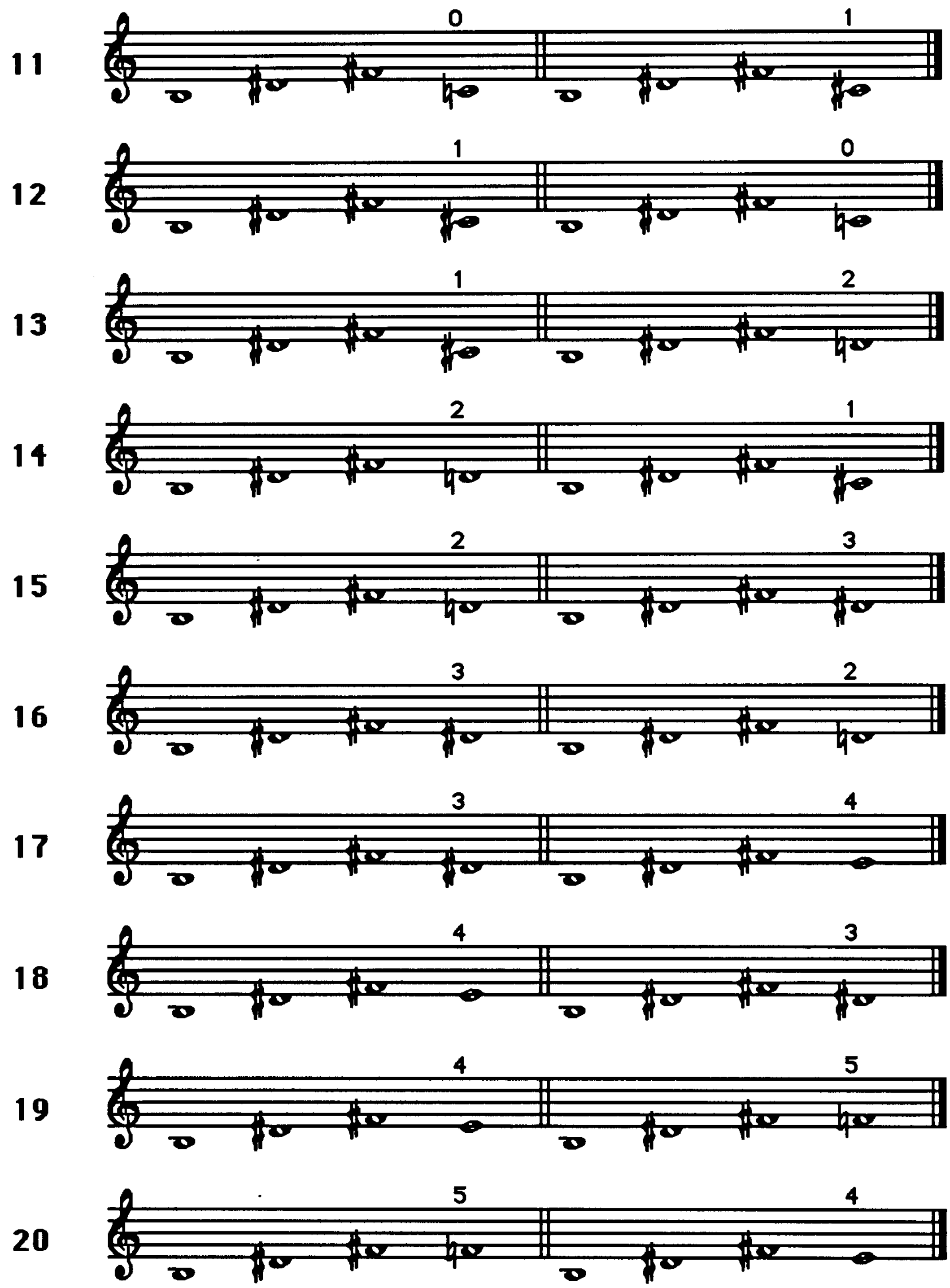
Experiment Two Materials (Chapter 6)

Music Notation

The following is a musical representation of the pitches presented for each of the twenty trials by the computer program in Appendix VII.

Numbers above the notes relate to the numbers in the data statements in the computer program. These numbers represent semitone offsets from the base note which is middle C (e.g. C=0, D flat=1, D=2, etc.).


1 0 0
2 1 1
3 1 1
4 2 2
5 2 2
6 3 3
7 3 3
8 4 4
9 4 4
10 5 5


11 0 1
12 1 0
13 1 2
14 2 1
15 2 3
16 3 2
17 3 4
18 4 3
19 4 5
20 5 4

APPENDIX VII

Experiment Two Materials (Chapter 6)

BBC Microcomputer Program

The following is a BASIC listing of the computer program which was written to present the experimental materials and record reaction time responses.

The musical pitches of the twenty trials which it presents are notated in Appendix VI.

The program may be found on the accompanying diskette as "$.MB2".

```
10 *KEY0RUN|M
20 *KEY1LIST|M
30 *KEY9SAVE"MB2"|M
40 :
50 ON ERROR GOTO 3160
60 :
70 REM ***** REACTION TIME TEST version 27.2.91 *****
80 REM ***** Triad context plus suffix *****
90 REM ***** BBC Microcomputer *****
100 REM ***** by Richard E. Hodges *****
110 :
120 REM ***** USER INPUT *****
130 REM ***** DATE *****
140 DATE$="28.2.91"
150 REM ***** TEST DESCRIPTOR *****
160 TEST$="TEST TWO"
170 REM ***** NUMBER OF QUESTIONS *****
180 Q=20
190 REM ***** PAUSE TIME FOR INSTRUCTIONS *****
200 REM ***** SHOULD BE 199 *****
210 PAUSE=199
220 :
230 REM ***** PROCEDURE CALLS *****
240 MODE 7
250 PROCinitialise
260 PROCtitle
270 PROCrandom
280 PROCgetname
290 PROCinstructions
```

```
300 PROCtestloop
310 PROCthankyou
320 PROCwritedata
330 PROCtitle
340 PROCtestend
350 PROCreset
360 END
370 :
380 :
390 REM ***** TITLE PAGE *****
400 DEFPROCtitle
410 CLS
420 REM ***** TURN OFF CURSOR *****
430  VDU 23,1,0,0,0,0,0,0,0
440 FOR I=1 TO 106
450 PROCrandomcolour
460 PRINT CHR$(COL);"MUSIC TEST   ";
470 NTE=40+I/2
480 PROCnoteon
490 FOR D=1 TO RND(30)+10
500 NEXT D
510 PROCnoteoff
520 NEXT I
530 ENDPROC
540 :
550 REM ***** INITIALISE *****
560 DEFPROCinitialise
570 REM ***** SET MIDI INTERFACE *****
580 ?&FC00=&3:?&FC00=&15
590 :
600 REM ***** DISABLE AUTO REPEAT DELAY ON KEYBOARD *****
610 *FX 11,0
620 :
630 REM ***** A TO D CONVERTER OFF *****
640 *FX 16,0
650 :
660 REM ***** TURN OFF CURSOR *****
670  VDU 23,1,0,0,0,0,0,0,0
680 :
690 REM ***** SET COLOURS *****
700 RED$=CHR$(129)
710 GREEN$=CHR$(130)
720 YELLOW$=CHR$(131)
730 BLUE$=CHR$(132)
740 MAGENTA$=CHR$(133)
750 CYAN$=CHR$(134)
760 WHITE$=CHR$(135)
770 DH$=CHR$(141)
780 :
790 REM ***** DIMENSION ARRAYS *****
800 DIM A(Q):DIM B(Q):DIM N(Q)
810 DIM DELAY(Q): DIM RESPONSE$(Q)
820 FOR I=1 TO Q
830 READ A(I):READ B(I)
840 NEXT I
850 ENDPROC
860 :
870 REM ***** RANDOM COLOUR *****
880 DEFPROCrandomcolour
```

```
890 COL=128+RND(7)
900 ENDPROC
910 :
920 REM ***** GENERATE RANDOM NUMBERS *****
930 DEFPROCrandom
940 FOR G=1 TO Q
950 R=INT(RND(1)*Q+1)
960 F%=0
970 FOR T=1 TO G
980 IF N(T)=R THEN F%=1
990 NEXT T
1000 N(G)=R
1010 IF F%=1 THEN GOTO 950
1020 NEXT G
1030 ENDPROC
1040 :
1050 REM ***** GET NAME FOR RESPONSES *****
1060 DEFPROCgetname
1070 CLS
1080 DOWN=7:TEXT$=CYAN$+"WHAT IS YOUR NAME ?"
1090 PROCcentretext
1100 A=0:NAME$=""
1110 REPEAT
1120 G=GET:A=A+1
1130 NAME$=NAME$+CHR$(G)
1140 DOWN=11:TEXT$=DH$+GREEN$+NAME$+"  "
1150 PROCcentretext
1160 UNTIL G=13 OR G=127
1170 IF A<2 OR G=127 THEN 1070
1180 CLS
1190 DOWN=14:TEXT$=" Pleased to meet you, "+NAME$
1200 PROCcentretext
1210 PROCpause
1220 PROCpressspace
1230 ENDPROC
1240 :
1250 REM ***** INSTRUCTIONS *****
1260 DEFPROCinstructions
1270 CLS
1280 DOWN=9:TEXT$=WHITE$+"YOU WILL
HEAR"+GREEN$+"TWO"+WHITE$+"4-NOTE TUNES"
1290 PROCcentretext
1300 PROCpause
1310 PROCpressspace
1320 DOWN=9:TEXT$=RED$+"THEY MAY BE THE "+YELLOW$+"SAME"
1330 PROCcentretext
1340 PROCpause
1350 PROCpressspace
1360 DOWN=9:TEXT$=RED$+"OR THEY MAY BE "+BLUE$+"DIFFERENT"
1370 PROCcentretext
1380 PROCpause
1390 PROCpressspace
1400 DOWN=5:TEXT$="PRESS "+YELLOW$+"'S'"+WHITE$+" IF THEY ARE
THE SAME"
1410 PROCcentretext
1420 PROCpause:PROCpause
1430 DOWN=9:TEXT$="PRESS "+BLUE$+"'D'"+WHITE$+" IF THEY ARE
DIFFERENT"
1440 PROCcentretext
```

```
1450 PROCpause
1460 DOWN=20:TEXT$=MAGENTA$+"PRESS SPACE-BAR TO BEGIN"
1470 PROCcentretext
1480 REM ***** FLUSH KEYBOARD BUFFER *****
1490 *FX 21,0
1500 G=0:REPEAT:G=GET:UNTIL G=32
1510 CLS:PROCpause
1520 ENDPROC
1530 :
1540 REM ***** CENTRE TEXT *****
1550 DEFPROCcentretext
1560 IF LEN(TEXT$)/2<>LEN(TEXT$+" ")/2 THEN TEXT$=TEXT$+" "
1570 PRINT TAB(19-(LEN(TEXT$)/2),DOWN),DH$+TEXT$
1580 PRINT TAB(19-(LEN(TEXT$)/2),DOWN+1),DH$+TEXT$
1590 ENDPROC
1600 :
1610 REM ***** PAUSE *****
1620 DEFPROCpause
1630 T=TIME:REPEAT UNTIL TIME>T+PAUSE
1640 ENDPROC
1650 :
1660 REM ***** TEST LOOP *****
1670 DEFPROCtestloop
1680 FOR L=1 TO Q
1690 I=N(L)
1700 REM ***** 60 = Midi Note Number for Middle C *****
1710 BASE=60
1720 CLS
1730 PROCrandomcolour
1740 DOWN=3:TEXT$=CHR$(COL)+"Test Item Number "+STR$(L)
1750 PROCcentretext
1760 PROCpause
1770 PROCplaytriad
1780 T2=TIME
1790 NTE=A(I)+BASE
1800 PROCnoteon
1810 PROCdelay
1820 PROCnoteoff
1830 REPEAT UNTIL TIME>T2+299
1840 :
1850 PROCplaytriad
1860 NTE=B(I)+BASE
1870 PROCnoteon
1880 PROCdelay1
1890 REPEAT UNTIL TIME>T1+99
1900 PROCnoteoff
1910 IF B<1 OR B>2 THEN PROCtestbutton:PROCprintresponse
1920 PROCpressspace
1930 NEXT L
1940 ENDPROC
1950 :
1960 DEFPROCplaytriad
1970 NTE=59:PROCplay
1980 NTE=63:PROCplay
1990 NTE=66:PROCplay
2000 ENDPROC
2010 :
2020 DEFPROCplay
2030 PROCnoteon
```

```
2040 T=TIME:REPEAT UNTIL TIME>T+75
2050 PROCnoteoff
2060 ENDPROC
2070 :
2080 REM ***** TEST BUTTON *****
2090 DEFPROCtestbutton
2100 REPEAT
2110 B=ADVAL(0) AND 3
2120 UNTIL B=1 OR B=2
2130 DELAY(I)=TIME-T1
2140 ENDPROC
2150 :
2160 REM ***** PRINT REPONSE *****
2170 DEFPROCprintresponse
2180 TEXT$="Delay is "+STR$(DELAY(I))+" centi-seconds"
2190 REM PRINT TAB(20-LEN(TEXT$)/2,18);TEXT$
2200 IF B=1 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+BLUE$+"DIFFERENT":PROCcentretext
2210 IF B=2 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+YELLOW$+"SAME":PROCcentretext
2220 IF B=1 THEN RESPONSE$(I)="DIFFERENT"
2230 IF B=2 THEN RESPONSE$(I)="SAME"
2240 ENDPROC
2250 :
2260 REM ***** DATA STATEMENTS *****
2270 REM ***** SAME STIMULI *****
2280 DATA 0, 0
2290 DATA 1, 1
2300 DATA 1, 1
2310 DATA 2, 2
2320 DATA 2, 2
2330 DATA 3, 3
2340 DATA 3, 3
2350 DATA 4, 4
2360 DATA 4, 4
2370 DATA 5, 5
2380 REM ***** DIFFERENT STIMULI *****
2390 DATA 0, 1
2400 DATA 1, 0
2410 DATA 1, 2
2420 DATA 2, 1
2430 DATA 2, 3
2440 DATA 3, 2
2450 DATA 3, 4
2460 DATA 4, 3
2470 DATA 4, 5
2480 DATA 5, 4
2490 :
2500 REM ***** PRESS SPACE FOR NEXT ITEM *****
2510 DEFPROCpressspace
2520 T=TIME:REPEAT UNTIL TIME>T+99
2530 DOWN=20:TEXT$=CYAN$+"PRESS SPACE BAR TO CONTINUE"
2540 PROCcentretext
2550 REM ***** FLUSH KEYBOARD BUFFER *****
2560 *FX 21,0
2570 G=0:REPEAT:G=GET:UNTIL G=32
2580 CLS:T=TIME:REPEAT UNTIL TIME>T+199
2590 ENDPROC
2600 :
```

```
2610 REM ***** MIDI NOTE ON *****
2620 DEFPROCnoteon
2630 ?&FC01=&90:?&FC01=NTE:?&FC01=64
2640 ENDPROC
2650 :
2660 REM ***** MIDI NOTE OFF *****
2670 DEFPROCnoteoff
2680 ?&FC01=&80:?&FC01=NTE:?&FC01=0
2690 ENDPROC
2700 :
2710 REM ***** DELAY *****
2720 DEFPROCdelay
2730 T=TIME:REPEAT UNTIL TIME>T+99
2740 ENDPROC
2750 :
2760 REM ***** DELAY1 *****
2770 DEFPROCdelay1
2780 T1=TIME
2790 REPEAT
2800 B=ADVAL(0) AND 3
2810 UNTIL B=1 OR B=2 OR TIME>T1+99
2820 IF B=1 OR B=2 THEN DELAY(I)=TIME-T1:PROCprintresponse
2830 IF B=3 THEN DOWN=7:TEXT$="DON'T PRESS BOTH BUTTONS
TOGETHER":PROCcentretext
2840 ENDPROC
2850 :
2860 REM ***** THANK YOU *****
2870 DEFPROCthankyou
2880 CLS
2890 DOWN=9:TEXT$=MAGENTA$+"THANK YOU FOR YOUR HELP"
2900 PROCcentretext
2910 ENDPROC
2920 :
2930 REM ***** WRITE DATA TO FILE *****
2940 DEFPROCwritedata
2950 RESULTS=OPENOUT("R."+NAME$)
2960 PRINT #RESULTS,DATE$,TEST$,Q
2970 FOR I=1 TO Q
2980 PRINT #RESULTS,N(I)
2990 NEXT I
3000 FOR I=1 TO Q
3010 PRINT #RESULTS,A(I),B(I)
3020 PRINT #RESULTS,DELAY(I),RESPONSE$(I)
3030 NEXT I
3040 CLOSE #RESULTS
3050 PROCpause
3060 ENDPROC
3070 :
3080 REM ***** TEST ENDING *****
3090 DEFPROCtestend
3100 REM ***** FLUSH KEYBOARD BUFFER *****
3110 *FX 21,0
3120 G=0:REPEAT:G=GET:UNTIL G<>0
3130 CLS
3140 ENDPROC
3150 :
3160 REM ***** ERROR HANDLING ROUTINE *****
3170 CLS: REPORT
3180 PRINT " AT LINE ";ERL
```

```
3190 PROCreset
3200 END
3210 :
3220 REM ***** RESET KEYBOARD AUTO-REPEAT RATE *****
3230 DEFPROCreset
3240 *FX 12,0
3250 ENDPROC
3260 :
3270 REM ***** TEST RANDOM NUMBERS *****
3280 FOR I=1 TO Q
3290 PRINT N(I)
3300 NEXT I
3310 END
3320 :
3330 REM ***** TEST NOTES IN ARRAY *****
3340 FOR I=1 TO Q
3350 PRINT A(I),B(I)
3360 NEXT I
3370 END
```

The following is an example of one of the BASIC computer programs which computed means and standard deviations directly from the computer files created by the previous program. The program computes the means and standard deviations of the correct only responses of all subjects in all years.

The program may be found on the accompanying diskette as "$.SDALL".

```
 10 *KEY9SAVE"$.SDALL"|M
 20 REM ***** READ RT FROM DISK *****
 30 REM ***** Version 11.4.91 *****
 40 REM ***** to compute all years correct only *****
 50 MODE 3
 60 REM ***** PRINTER LINE FEED *****
 70 *FX 6,0
 80 *DIR R
 90 Q=20
100 REM ***** MAXFILE = 33 *****
110 MAXFILE=33
120 :
130 DIM A(Q):DIM B(Q)
140 DIM EXAM(40)
150 DIM FLAG(Q)
160 DIM DELAY(Q)
170 DIM RESPONSE$(Q)
180 DIM NUM(Q)
190 DIM GT(Q)
200 DIM MA(MAXFILE,Q)
210 DIM OBS(Q)
220 DIM SQUARE(Q)
230 DIM MEAN(Q)
240 DIM CORRTERM(Q)
250 DIM SUM(Q)
260 DIM VAR(Q)
270 DIM N(Q)
280 DIM T(MAXFILE)
290 DIM D(MAXFILE)
300 :
310 REM ***** LOOP TO READ IN FILES *****
320 FOR FILES=1 TO MAXFILE
330 READ D
340 READ NAME$
350 PRINT NAME$;"      ";
360 :
370 REM ***** YEAR 6 *****
380 DATA 0
390 DATA STEPHEN
400 DATA 0
410 DATA JONTY
420 DATA 0
430 DATA MATHEW
440 DATA 0
450 DATA NIGEL-B
```

```
460 DATA 0
470 DATA MARTIN
480 DATA 0
490 DATA LIZZY
500 REM ***** YEAR 5 *****
510 DATA 0
520 DATA ROBERT
530 DATA 0
540 DATA BURDY
550 DATA 0
560 DATA LISA
570 DATA 0
580 DATA LAURA
590 DATA 0
600 DATA NICKY
610 DATA 0
620 DATA KATIE
630 DATA 0
640 DATA ALICE
650 REM ***** YEAR 4 *****
660 DATA 0
670 DATA CHRIS
680 DATA 0
690 DATA PETER
700 DATA 0
710 DATA JENNY
720 DATA 0
730 DATA CATH
740 DATA 0
750 DATA EMMA
760 REM ***** YEAR 3 *****
770 DATA 2
780 DATA STEP-M
790 DATA 2
800 DATA ADAM
810 DATA 2
820 DATA ALEX-D
830 DATA 2
840 DATA NIGEL-P
850 DATA 2
860 DATA OLIVER
870 DATA 2
880 DATA REBECCA
890 DATA 2
900 DATA LOUISE
910 DATA 2
920 DATA KRISTY
930 REM DATA 0
940 REM DATA HELEN
950 REM ***** YEAR 2 *****
960 DATA 2
970 DATA TONY
980 DATA 2
990 DATA WILLIAM
1000 DATA 2
1010 DATA PAUL
1020 DATA 2
1030 DATA GEORGE
1040 DATA 2
```

```
1050 DATA GEMMA
1060 DATA 2
1070 DATA STEPH
1080 DATA 2
1090 DATA JOH_
1100 :
1110 IF D=0 THEN *DRIVE 0
1120 IF D=2 THEN *DRIVE 2
1130 FILENO=OPENIN(NAME$)
1140 INPUT #FILENO,DATE$,TEST$,Q
1150 PRINT "Date is ";DATE$;"   ";
1160 PRINT "Test is ";TEST$;"   ";
1170 PRINT "No. of Questions is ";Q
1180 FOR I=1 TO Q
1190 INPUT #FILENO,NUM(I)
1200 NEXT I
1210 :
1220 T1=0:DELAY1=0
1230 T2=0:DELAY2=0
1240 FOR I=1 TO Q
1250 INPUT #FILENO,A(I),B(I),DELAY(I),RESPONSE$(I)
1260 MA(FILES,I)=DELAY(I)
1270 EXAM(I)=EXAM(I)+DELAY(I)
1280 A$=RESPONSE$(I)
1290 PRINT I;
1300 PRINT "        ";
1310 REM PRINT A(I),B(I),DELAY(I),,"        ";RESPONSE$(I),;
1320 PRINT DELAY(I);
1330 PRINT "        ";
1340 IF A(I)=B(I) AND A$="SAME" THEN T1=T1+1:PRINT " CORRECT";
1350 IF A(I)=B(I) AND A$="SAME" THEN DELAY1=DELAY1+DELAY(I)
1360 IF A(I)=B(I) AND A$="DIFFERENT" THEN PRINT "   WRONG";
1370 IF A(I)=B(I) AND A$="DIFFERENT" THEN EXAM(I)=EXAM(I)-DELAY(I)
1380 IF A(I)=B(I) AND A$="DIFFERENT" THEN MA(FILES,I)=0
1390 IF A(I)=B(I) AND A$="DIFFERENT" THEN FLAG(I)=FLAG(I)+1
1400 IF A(I)<>B(I) AND A$="DIFFERENT" THEN T2=T2+1:PRINT " CORRECT";
1410 IF A(I)<>B(I) AND A$="DIFFERENT" THEN DELAY2=DELAY2+DELAY(I)
1420 IF A(I)<>B(I) AND A$="SAME" THEN PRINT "   WRONG";
1430 IF A(I)<>B(I) AND A$="SAME" THEN EXAM(I)=EXAM(I)-DELAY(I)
1440 IF A(I)<>B(I) AND A$="SAME" THEN MA(FILES,I)=0
1450 IF A(I)<>B(I) AND A$="SAME" THEN FLAG(I)=FLAG(I)+1
1460 PRINT EXAM(I);
1470 PRINT
1480 NEXT I
1490 PRINT
1500 IF T1=0 OR DELAY1=0 THEN GOTO 1530
1510 PRINT "MEAN CORRECT SAME = ";DELAY1/T1
1520 PRINT "Percentage Correct = ";(T1/10)*100
1530 IF T2=0 OR DELAY2=0 THEN GOTO 1570
1540 PRINT "MEAN CORRECT DIFFERENT = ";DELAY2/T2
1550 PRINT "Percentage Correct = ";(T2/10)*100
1560 :
1570 CLOSE #FILENO
1580 PRINT
1590 PRINT "Same      correct = ";T1
1600 PRINT "Different correct = ";T2
1610 PRINT
1620 F=F+T1
1630 G=G+T2
```

```
1640 PRINT "Total correct same      = ";F
1650 PRINT "Total correct different = ";G
1660 PRINT:PRINT
1670 NEXT FILES
1680 :
1690 FOR I=1 TO 20
1700 PRINT FLAG(I)
1710 WRONG=WRONG+FLAG(I)
1720 NEXT
1730 PRINT
1740 PRINT WRONG; " Mistakes"
1750 PRINT
1760 :
1770 REM ***** PRINT MEANS *****
1780 PRINT "MEANS"
1790 FOR A=1 TO 20
1800 IF A=11 THEN PRINT:REPEAT:PRESS=GET:UNTIL PRESS=32
1810 GT(A)=EXAM(A)
1820 PRINT "TRIAL ";A;"      ";GT(A)/((FILES-1)-FLAG(A))
1830 PRINT "TOTAL ";GT(A);" / ";"FILES-1 (";FILES-1;") - FLAG(WRONG)
";FLAG(A)
1840 NEXT A
1850 :
1860 PRINT
1870 FOR A=1 TO 10:L=L+GT(A):NEXT A
1880 PRINT "MEAN SAME      = ";L;" / ";F," = ";
1890 PRINT L/F
1900 FOR A=11 TO 20:M=M+GT(A):NEXT A
1910 PRINT "MEAN DIFFERENT = ";M;" / ";G," = ";
1920 PRINT M/G
1930 :
1940 FOR FILE=1 TO MAXFILE
1950 FOR TRIAL=11 TO 20
1960 T(FILE)=T(FILE)+MA(FILE,TRIAL)
1970 IF MA(FILE,TRIAL)=0 THEN D(FILE)=D(FILE)+1
1980 NEXT TRIAL
1990 PRINT
2000 PRINT T(FILE);" is the total different for subject ";FILE;
2010 PRINT D(FILE); " errors"
2011 IF D(FILE)=10 THEN GOTO 2050
2020 PRINT T(FILE)/(10-D(FILE))
2030 MAX=MAX+T(FILE)
2040 MAX1=MAX1+(T(FILE)/(10-D(FILE)))
2050 NEXT FILE
2060 PRINT
2070 PRINT MAX;" is the total for all subjects"
2080 PRINT MAX1/(MAXFILE);" is the mean"
2090 REPEAT UNTIL GET=32
2100 :
2110 REM ***** PRINTOUT *****
2120 FOR I=1 TO Q
2130 PRINT:PRINT " TRIAL ";I
2140 FOR FILE=1 TO MAXFILE
2150 PRINT MA(FILE,I);
2160 NEXT FILE
2170 NEXT I
2180 :
2190 REM ***** Calculate Standard Deviation *****
2200 PRINT
```

```
2210 FOR I=1 TO Q
2220 FOR FILE=1 TO MAXFILE
2230 OBS(I)=OBS(I)+MA(FILE,I)
2240 IF MA(FILE,I)<>0 THEN N(I)=N(I)+1
2250 SQUARE(I)=SQUARE(I)+(MA(FILE,I)^2)
2260 NEXT FILE
2270 :
2280 PRINT
2290 CORRTERM(I)=(OBS(I)^2)/N(I)
2300 PRINT "Correction Term = ";CORRTERM(I);" = ";(OBS(I)^2);" / ";N(I)
2310 SUM(I)=SQUARE(I)-CORRTERM(I)
2320 PRINT "Sum of Squares  = ";SUM(I);" = ";SQUARE(I);" - ";CORRTERM(I)
2330 VAR(I)=SUM(I)/(N(I)-1)
2340 PRINT "Variance        = ";VAR(I);" = ";SUM(I);" / ";(N(I)-1)
2350 NEXT I
2360 :
2370 PRINT
2380 FOR I=1 TO 20
2390 PRINT "Variance = ";VAR(I);
2400 PRINT TAB(40);"S.D. = ";SQR(VAR(I))
2410 IF I=10 THEN PRINT:REPEAT UNTIL GET=32
2420 NEXT I
```

APPENDIX VIII

Experiment Three Materials (Chapter 7)

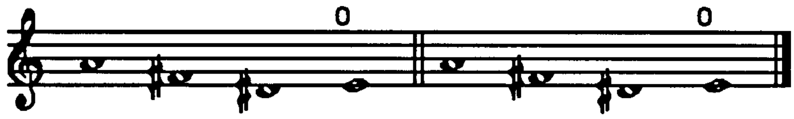
Music Notation

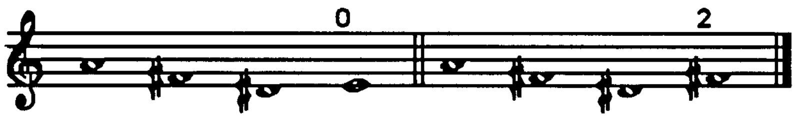
The following is a musical representation of the pitches presented for each of the thirty-six trials by the computer program in Appendix IX.

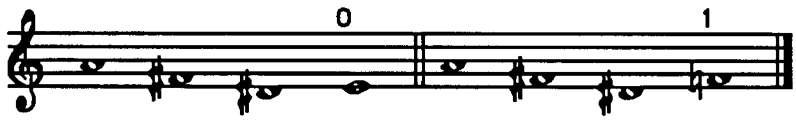
Numbers above the notes relate to the numbers in the data statements in the computer program. These numbers represent semitone offsets from the base note which is E (i.e. E=0, F=1, F sharp=2, etc.).

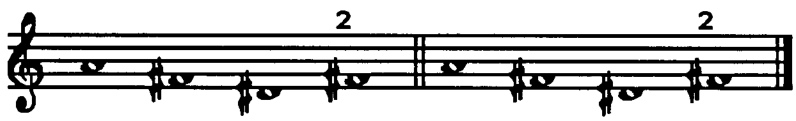
The three letter code after the trial number is coded as follows:

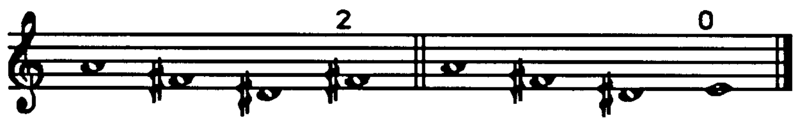
S=Same
D=Different
N=Narrow (spread in relation to circle of fifths)
W=Wide (spread in relation to circle of fifths)

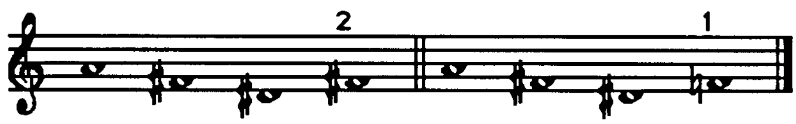
Trial 1 (SNN)


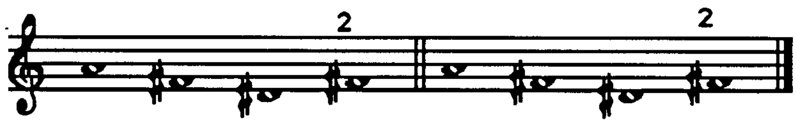
Trial 2 (DNN)


Trial 3 (DNW)


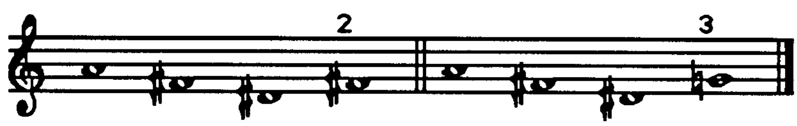
Trial 4 (SNN)


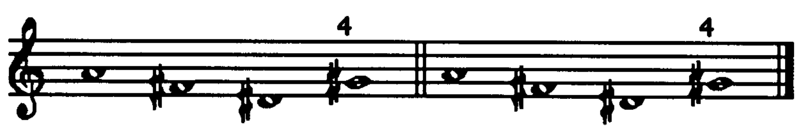
Trial 5 (DNN)


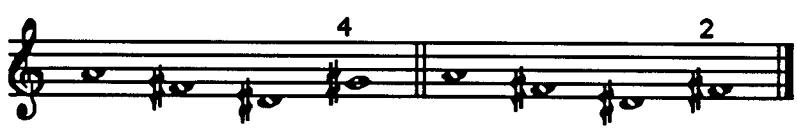
Trial 6 (DNW)


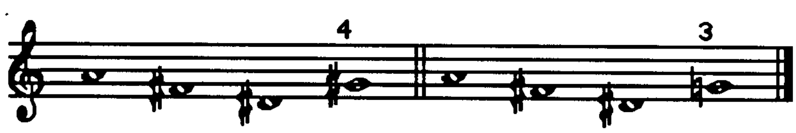
Trial 7 (SNN)


Trial 8 (DNN)

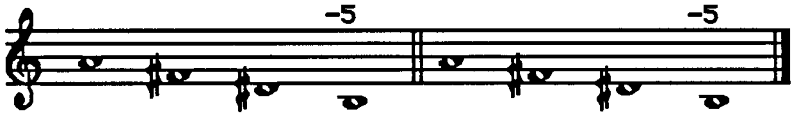

Trial 9 (DNW)


Trial 10 (SNN)


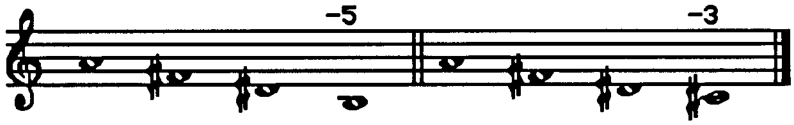
Trial 11 (DNN)


Trial 12 (DNW)


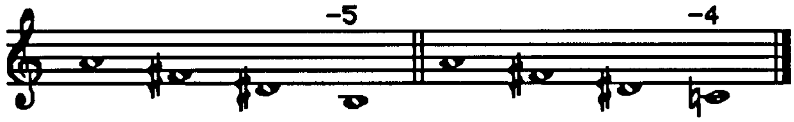
Trial 13 (SNN)



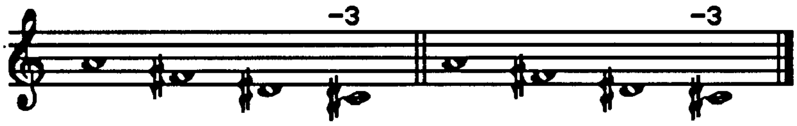
Trial 14 (DNN)



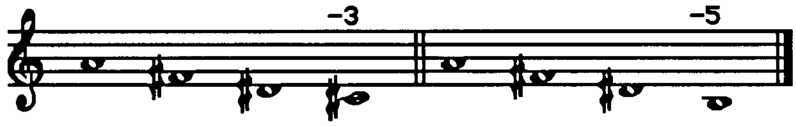
Trial 15 (DNW)



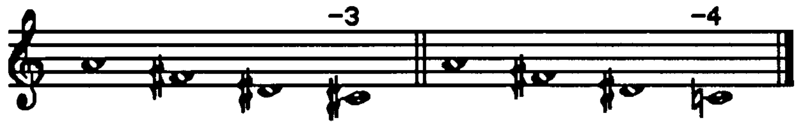
Trial 16 (SNN)



Trial 17 (DNN)



Trial 18 (DNW)

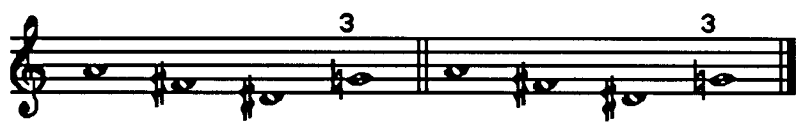


Trial 19 (SWW)

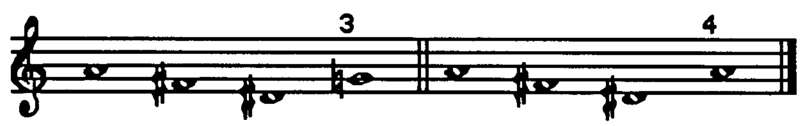
Trial 20 (DWN)

Trial 21 (DWW)

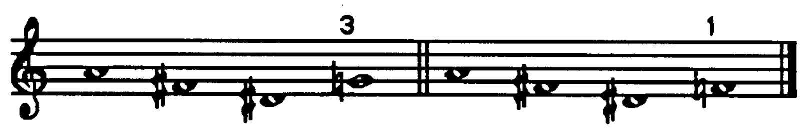
Trial 22 (SWW)



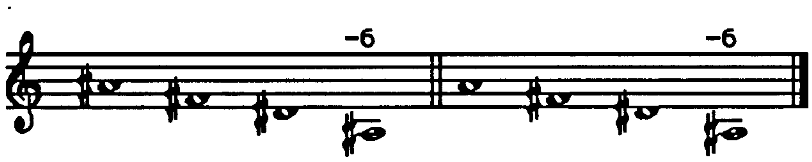
Trial 23 (DWN)



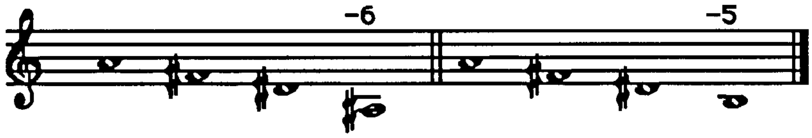
Trial 24 (DWW)



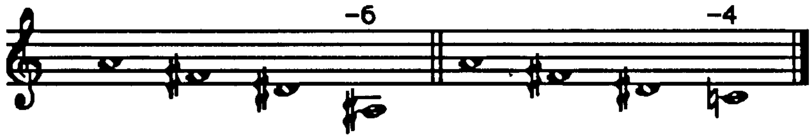
Trial 25 (SWW)



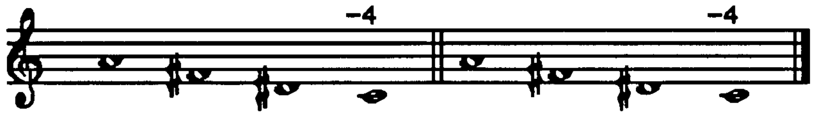
Trial 26 (DWN)



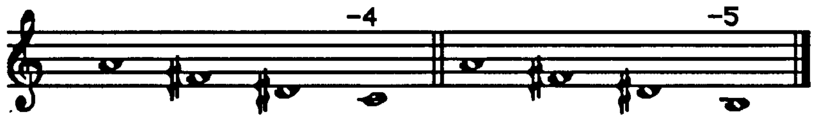
Trial 27 (DWW)



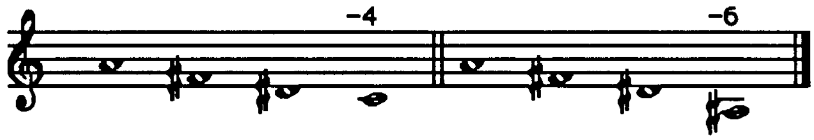
Trial 28 (SWW)



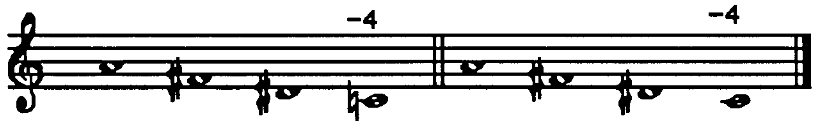
Trial 29 (DWN)



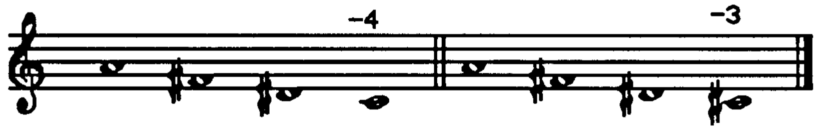
Trial 30 (DWW)



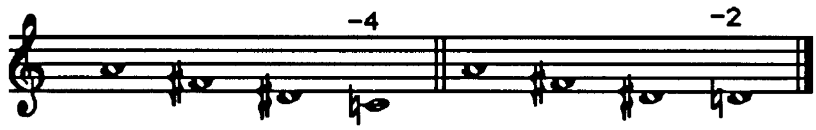
Trial 31 (SWW)



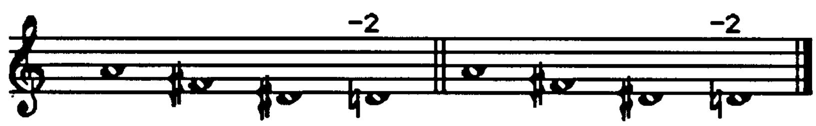
Trial 32 (DWN)



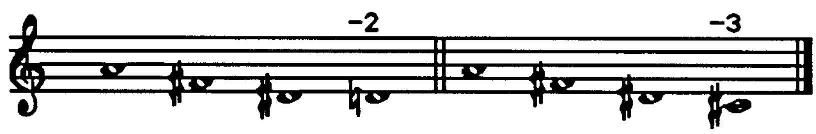
Trial 33 (DWW)



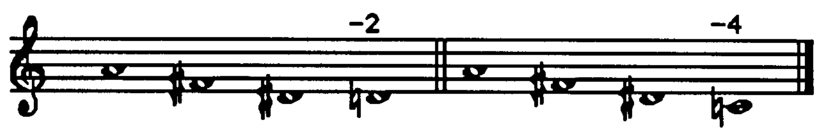
Trial 34 (SWW)



Trial 35 (DWN)



Trial 36 (DWW)



APPENDIX IX

Experiment Three Materials (Chapter 7)

BBC Microcomputer Program

The following is a BASIC listing of the computer program which was written to present the experimental materials and record reaction time responses.

The musical pitches of the thirty-six trials which it presents are notated in Appendix VIII.

The program may be found on the accompanying diskette as "$.MB3".

```
10 *KEY0RUN|M
20 *KEY1LIST|M
30 *KEY9SAVE"MB3"|M
40 :
50 ON ERROR GOTO 3460
60 :
70 REM ***** REACTION TIME TEST version 12.3.92 *****
80 REM ***** Diminished Triad context plus suffix *****
90 REM ***** BBC Microcomputer *****
100 REM ***** by Richard E. Hodges *****
110 :
120 REM ***** USER INPUT *****
130 REM ***** DATE *****
140 DATE$="10.3.92"
150 REM ***** TEST DESCRIPTOR *****
160 TEST$="TEST THREE"
170 REM ***** NUMBER OF QUESTIONS *****
180 Q=36
190 REM ***** PAUSE TIME FOR INSTRUCTIONS *****
200 REM ***** SHOULD BE 199 *****
210 PAUSE=199
220 :
230 REM ***** PROCEDURE CALLS *****
240 MODE 7
250 PROCinitialise
260 PROCtitle
270 PROCrandom
280 PROCgetname
```

```
290 PROCinstructions
300 PROCtestloop
310 PROCthankyou
320 PROCwritedata
330 PROCtitle
340 PROCtestend
350 PROCreset
360 END
370 :
380 :
390 REM ***** TITLE PAGE *****
400 DEFPROCtitle
410 CLS
420 REM ***** TURN OFF CURSOR *****
430 VDU 23,1,0,0,0,0,0,0,0
440 FOR I=1 TO 106
450 PROCrandomcolour
460 PRINT CHR$(COL);"MUSIC TEST   ";
470 NTE=40+I/2
480 PROCnoteon
490 FOR D=1 TO RND(30)+10
500 NEXT D
510 PROCnoteoff
520 NEXT I
530 ENDPROC
540 :
550 REM ***** INITIALISE *****
560 DEFPROCinitialise
570 REM ***** SET MIDI INTERFACE *****
580 ?&FC00=&3:?&FC00=&15
590 :
600 REM ***** DISABLE AUTO REPEAT DELAY ON KEYBOARD *****
610 *FX 11,0
620 :
630 REM ***** A TO D CONVERTER OFF *****
640 *FX 16,0
650 :
660 REM ***** TURN OFF CURSOR *****
670 VDU 23,1,0,0,0,0,0,0,0
680 :
690 REM ***** SET COLOURS *****
700 RED$=CHR$(129)
710 GREEN$=CHR$(130)
720 YELLOW$=CHR$(131)
730 BLUE$=CHR$(132)
740 MAGENTA$=CHR$(133)
750 CYAN$=CHR$(134)
760 WHITE$=CHR$(135)
770 DH$=CHR$(141)
780 :
790 REM ***** DIMENSION ARRAYS *****
800 DIM A(Q):DIM B(Q):DIM N(Q)
810 DIM DELAY(Q): DIM RESPONSE$(Q)
820 FOR I=1 TO Q
830 READ A(I):READ B(I)
840 NEXT I
850 ENDPROC
860 :
870 REM ***** RANDOM COLOUR *****
```

```
880 DEFPROCrandomcolour
890 COL=128+RND(7)
900 ENDPROC
910 :
920 REM ***** GENERATE RANDOM NUMBERS *****
930 DEFPROCrandom
940 FOR G=1 TO Q
950 R=INT(RND(1)*Q+1)
960 F%=0
970 FOR T=1 TO G
980 IF N(T)=R THEN F%=1
990 NEXT T
1000 N(G)=R
1010 IF F%=1 THEN GOTO 950
1020 NEXT G
1030 ENDPROC
1040 :
1050 REM ***** GET NAME FOR RESPONSES *****
1060 DEFPROCgetname
1070 CLS
1080 DOWN=7:TEXT$=CYAN$+"What is your name ?"
1090 PROCcentretext
1100 A=0:NAME$=""
1110 REPEAT
1120 G=GET:A=A+1
1130 NAME$=NAME$+CHR$(G)
1140 DOWN=11:TEXT$=DH$+GREEN$+NAME$+"  "
1150 PROCcentretext
1160 UNTIL G=13 OR G=127
1170 IF A<2 OR G=127 THEN 1070
1180 CLS
1190 DOWN=14:TEXT$=" Pleased to meet you, "+NAME$
1200 PROCcentretext
1210 PROCpause
1220 PROCpressspace
1230 ENDPROC
1240 :
1250 REM ***** INSTRUCTIONS *****
1260 DEFPROCinstructions
1270 CLS
1280 DOWN=9:TEXT$=WHITE$+"You will hear"+GREEN$+"Two"+WHITE$+"4-
note tunes"
1290 PROCcentretext
1300 PROCpause
1310 PROCpressspace
1320 DOWN=9:TEXT$=RED$+"They may be the "+YELLOW$+"SAME"
1330 PROCcentretext
1340 PROCpause
1350 PROCpressspace
1360 DOWN=9:TEXT$=RED$+"Or they may be "+BLUE$+"DIFFERENT"
1370 PROCcentretext
1380 PROCpause
1390 PROCpressspace
1400 DOWN=5:TEXT$="Press "+YELLOW$+"'S'"+WHITE$+" if they are the same"
1410 PROCcentretext
1420 PROCpause:PROCpause
1430 DOWN=9:TEXT$="Press "+BLUE$+"'D'"+WHITE$+" if they are different"
1440 PROCcentretext
1450 PROCpause
```

```
1460 DOWN=20:TEXT$=MAGENTA$+"Press space-bar to begin"
1470 PROCcentretext
1480 REM ***** FLUSH KEYBOARD BUFFER *****
1490 *FX 21,0
1500 G=0:REPEAT:G=GET:UNTIL G=32
1510 CLS:PROCpause
1520 ENDPROC
1530 :
1540 REM ***** CENTRE TEXT *****
1550 DEFPROCcentretext
1560 IF LEN(TEXT$)/2<>LEN(TEXT$+" ")/2 THEN TEXT$=TEXT$+" "
1570 PRINT TAB(19-(LEN(TEXT$)/2),DOWN),DH$+TEXT$
1580 PRINT TAB(19-(LEN(TEXT$)/2),DOWN+1),DH$+TEXT$
1590 ENDPROC
1600 :
1610 REM ***** PAUSE *****
1620 DEFPROCpause
1630 T=TIME:REPEAT UNTIL TIME>T+PAUSE
1640 ENDPROC
1650 :
1660 REM ***** TEST LOOP *****
1670 DEFPROCtestloop
1680 FOR L=1 TO Q
1690 I=N(L)
1700 REM ***** 60 = Midi Note Number for Middle C *****
1710 BASE=64
1720 CLS
1730 PROCrandomcolour
1740 DOWN=3:TEXT$=CHR$(COL)+"Test Trial Number "+STR$(L)
1750 PROCcentretext
1760 PROCpause
1770 PROCplaytriad
1780 T2=TIME
1790 NTE=A(I)+BASE
1800 PROCnoteon
1810 PROCdelay
1820 PROCnoteoff
1830 REPEAT UNTIL TIME>T2+299
1840 :
1850 PROCplaytriad
1860 NTE=B(I)+BASE
1870 PROCnoteon
1880 PROCdelay1
1890 REPEAT UNTIL TIME>T1+99
1900 PROCnoteoff
1910 IF B<1 OR B>2 THEN PROCtestbutton:PROCprintresponse
1920 PROCpressspace
1930 NEXT L
1940 ENDPROC
1950 :
1960 DEFPROCplaytriad
1970 NTE=69:PROCplay
1980 NTE=66:PROCplay
1990 NTE=63:PROCplay
2000 ENDPROC
2010 :
2020 DEFPROCplay
2030 PROCnoteon
2040 T=TIME:REPEAT UNTIL TIME>T+75
```

```
2050 PROCnoteoff
2060 ENDPROC
2070 :
2080 REM ***** TEST BUTTON *****
2090 DEFPROCtestbutton
2100 REPEAT
2110 B=ADVAL(0) AND 3
2120 UNTIL B=1 OR B=2
2130 DELAY(I)=TIME-T1
2140 ENDPROC
2150 :
2160 REM ***** PRINT REPONSE *****
2170 DEFPROCprintresponse
2180 TEXT$="Delay is "+STR$(DELAY(I))+" centi-seconds"
2190 REM PRINT TAB(20-LEN(TEXT$)/2,18);TEXT$
2200 IF B=1 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+BLUE$+"DIFFERENT":PROCcentretext
2210 IF B=2 THEN DOWN=9:TEXT$=WHITE$+"You
Pressed"+YELLOW$+"SAME":PROCcentretext
2220 IF B=1 THEN RESPONSE$(I)="DIFFERENT"
2230 IF B=2 THEN RESPONSE$(I)="SAME"
2240 ENDPROC
2250 :
2260 REM ***** DATA STATEMENTS *****
2270 REM ***** NARROW STIMULI *****
2280 :
2290 DATA  0, 0
2300 DATA  0, 2
2310 DATA  0, 1
2320 :
2330 DATA  2, 2
2340 DATA  2, 0
2350 DATA  2, 1
2360 :
2370 DATA  2, 2
2380 DATA  2, 4
2390 DATA  2, 3
2400 :
2410 DATA  4, 4
2420 DATA  4, 2
2430 DATA  4, 3
2440 :
2450 DATA -5,-5
2460 DATA -5,-3
2470 DATA -5,-4
2480 :
2490 DATA -3,-3
2500 DATA -3,-5
2510 DATA -3,-4
2520 :
2530 REM ***** WIDE STIMULI *****
2540 :
2550 DATA  1, 1
2560 DATA  1, 4
2570 DATA  1, 3
2580 :
2590 DATA  3, 3
2600 DATA  3, 4
2610 DATA  3, 1
```

```
2620 :
2630 DATA -6,-6
2640 DATA -6,-5
2650 DATA -6,-4
2660 :
2670 DATA -4,-4
2680 DATA -4,-5
2690 DATA -4,-6
2700 :
2710 DATA -4,-4
2720 DATA -4,-3
2730 DATA -4,-2
2740 :
2750 DATA -2,-2
2760 DATA -2,-3
2770 DATA -2,-4
2780 :
2790 :
2800 REM ***** PRESS SPACE FOR NEXT ITEM *****
2810 DEFPROCpressspace
2820 T=TIME:REPEAT UNTIL TIME>T+99
2830 DOWN=20:TEXT$=CYAN$+"Press space-bar to continue"
2840 PROCcentretext
2850 REM ***** FLUSH KEYBOARD BUFFER *****
2860 *FX 21,0
2870 G=0:REPEAT:G=GET:UNTIL G=32
2880 CLS:T=TIME:REPEAT UNTIL TIME>T+199
2890 ENDPROC
2900 :
2910 REM ***** MIDI NOTE ON *****
2920 DEFPROCnoteon
2930 ?&FC01=&90:?&FC01=NTE:?&FC01=64
2940 ENDPROC
2950 :
2960 REM ***** MIDI NOTE OFF *****
2970 DEFPROCnoteoff
2980 ?&FC01=&80:?&FC01=NTE:?&FC01=0
2990 ENDPROC
3000 :
3010 REM ***** DELAY *****
3020 DEFPROCdelay
3030 T=TIME:REPEAT UNTIL TIME>T+99
3040 ENDPROC
3050 :
3060 REM ***** DELAY1 *****
3070 DEFPROCdelay1
3080 T1=TIME
3090 REPEAT
3100 B=ADVAL(0) AND 3
3110 UNTIL B=1 OR B=2 OR TIME>T1+99
3120 IF B=1 OR B=2 THEN DELAY(I)=TIME-T1:PROCprintresponse
3130 IF B=3 THEN DOWN=7:TEXT$="DON'T PRESS BOTH BUTTONS
TOGETHER":PROCcentretext
3140 ENDPROC
3150 :
3160 REM ***** THANK YOU *****
3170 DEFPROCthankyou
3180 CLS
3190 DOWN=9:TEXT$=MAGENTA$+"THANK YOU FOR YOUR HELP"
```

```
3200 PROCcentretext
3210 ENDPROC
3220 :
3230 REM ***** WRITE DATA TO FILE *****
3240 DEFPROCwritedata
3250 RESULTS=OPENOUT("R."+NAME$)
3260 PRINT #RESULTS,DATE$,TEST$,Q
3270 FOR I=1 TO Q
3280 PRINT #RESULTS,N(I)
3290 NEXT I
3300 FOR I=1 TO Q
3310 PRINT #RESULTS,A(I),B(I)
3320 PRINT #RESULTS,DELAY(I),RESPONSE$(I)
3330 NEXT I
3340 CLOSE #RESULTS
3350 PROCpause
3360 ENDPROC
3370 :
3380 REM ***** TEST ENDING *****
3390 DEFPROCtestend
3400 REM ***** FLUSH KEYBOARD BUFFER *****
3410 *FX 21,0
3420 G=0:REPEAT:G=GET:UNTIL G<>0
3430 CLS
3440 ENDPROC
3450 :
3460 REM ***** ERROR HANDLING ROUTINE *****
3470 CLS: REPORT
3480 PRINT " AT LINE ";ERL
3490 PROCreset
3500 END
3510 :
3520 REM ***** RESET KEYBOARD AUTO-REPEAT RATE *****
3530 DEFPROCreset
3540 *FX 12,0
3550 ENDPROC
3560 :
3570 REM ***** TEST RANDOM NUMBERS *****
3580 FOR I=1 TO Q
3590 PRINT N(I)
3600 NEXT I
3610 END
3620 :
3630 REM ***** TEST NOTES IN ARRAY *****
3640 FOR I=1 TO Q
3650 PRINT A(I),B(I)
3660 NEXT I
3670 END
```

APPENDIX X

Experiment Three Baseline Reaction Time (Chapter 7)

BBC Microcomputer Program

The following is a BASIC listing of the computer program which was written to present twenty randomised beeps and record reaction time responses directly to disk.

The program may be found on the accompanying diskette as "$.RTSPACE".

```
10 *KEY0SAVE"RTSPACE"|M
20 Q=20
30 DIM DELAY(20)
40 MODE 7
50 PRINT
60 INPUT; "What is your name: ";NAME$
70 FOR I=1 TO Q
80 T=TIME:REPEAT UNTIL TIME>T+300
90 T=TIME:REPEAT UNTIL TIME>T+RND(500)
100 *FX21,0
110 PRINT CHR$(7)
120 G=0
130 T=TIME
140 REPEAT:G=GET:UNTIL G=32
150 DELAY(I)=TIME-T
160 REM PRINT DELAY(I)
170 NEXT I
180 :
190 RESULTS+OPENOUT("R."+NAME$)
200 FOR I=1 TO Q
210 PRINT #RESULTS, DELAY(I)
220 NEXT I
230 CLOSE #RESULTS
240 END
```