Cross-Modal and Synaesthetic Perception in Music and Vision

Lawrence D. Gerstley

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at
University of Stirling
1997

Supervisor: Prof. RJ Watt

Department of Psychology
Centre for Computational and Cognitive Neuroscience
University of Stirling
Scotland

ABSTRACT

This thesis is concerned with the cross-modal and synaesthetic perception of musical and visual stimuli. Each of these types of perception has been researched separately, and a hypothesis is presented here that accounts for both cross-modal matching and the development of synaesthesia. This hypothesis claims that sensory information can be evaluated in another modality by using a *scale of comparison* in that modality.

The first set of experiments examines normal subjects performing cross-modal matching with coloured circles and auditory stimuli that vary in complexity. It is shown that subjects use a variety of scales of comparison from both visual and auditory modalities to form matches. As the stimuli increase in complexity, the individual variation in cross-modal matching also increases.

The second set of experiments examines matching performance using higher order stimuli, by having subjects evaluate fragments of melodies and complete melodies on affective and descriptive adjective scales. Melodies were also matched with landscape scenes to examine if subjects could form matches between two highly complex sets of stimuli.

The final experiments examine synaesthetic associations with colour, evoked from music, letters, numbers, and other categorical information. Common features of synaesthesia from a population of synaesthetes are identified, and experiments performed to test the interference of the synaesthetic associations. Additional experiments are presented that explore the superior short-term memory of one synaesthete, and the role of his associations as a mnemonic device.

ACKNOWLEDGEMENTS

I would like to thank all of those folks from the CCCN that have provided examples of all types of behaviour that I have required in the course of researching and writing this thesis: especially the need to persevere throughout while never taking oneself too seriously. I am extremely grateful to Roger Watt for helping me achieve this Cadmeian victory, and his willingness to discuss minutiae even in Swedish sunshine. The residents of the Centre come and go, and often return once again—and now the Centre may be no more, but to me it will be a place where the same group dwells always, and I will always remember the help they have provided and the inspiration they provided in my life. Thanks to Trish Carlin for all the advice, help, and crying laughter; Roisin Ash for vocabulary lessons and stress relief, Paul Miller for the years of discussions; Ben Craven for helping to research virtually all problems; Fiona Biggam, Graham Dyson, Manoli Stamatakis, Ian Paterson, and Steve Dakin, for providing the optimism, and Will Goodall for reminding me that music is essentially about enjoyment as well as research.

Thanks to Shelley and Kirsty who were there through it all, and Rick, Kathy, Scott, Rich, Kate, Amanda, Karen, the Silvers, and the Siegels for anchoring me those things that are most important in life. Finally, I am most grateful to my family, to Rob, Julie, Mike, and Ellie, without whom I could never have begun, continued, or finished any of this. Their support inspires me every day. And to Rick—I *will* actually do it.

This thesis is dedicated to my parents, Michael and Eleanor Gerstley.

CONTENTS

1 Theories Of Cross-modal Perception and Synaesthesia

1.1 Introduction

A musical tone, a vivid colour, an acrid smell—each of these sensations is experienced in only a single human sense. However, no sensation is ever perceived in isolation. Aristotelian and Lockean primary qualities of sensations, such as number, size, magnitude, as well as other multiple sensory qualities, are an essential part of constructing representations of the world outside of the perceiver. The manner in which correspondences form between sensations, or common sensory characteristics are extracted from sensory experiences can reveal important information as to how personal experiences are integrated together to form these representations of the external world.

Certainly no sensory experience occurs in isolation. What might be considered as a purely auditory event occurs while the perceiver receives visual, tactile, propriaceptive, olfactory, and other sensory events. The process of learning which events provide salient information about the world to multiple modalities, and simultaneously learning which characteristics within experiences aid in discrimination and categorisation, are necessary parts of cognitive development.

This thesis will examine some of the cross-modal associations made between sensory events which occur within a mature cognitive system. That is, it examines what types of cross-modal matching normal individuals engage in when requested to examine sensory experiences in multiple modalities. In addition, synaesthesia and synaesthetic correspondences will be investigated, and the qualities of those unusual perceptions will be compared to cross-modal perception in normal individuals.

Synaesthesia is hypothesised to be a condition that arises from fixedness in cross-modal associations. In Cytowic's (1989) words, synaesthetes may be cognitive fossils, persevering in cross-modal perception processes that are left behind in the normal course of development. It is important to note, however, that Cytowic uses this term to imply that synaesthesia was earlier in man's evolution rather than earlier in an individual's development (Cytowic, 1989; pp. 21–22).

The thesis will also focus primarily on those cross-modal associations that include vision as one of the sensory modalities, usually when paired with audition. The assertion that vision is key among our senses may also mean that vision lends itself readily to a greater variety of sensory correspondences than any other sense. Also, a great majority of existing cross-modal research has been performed with vision as one of the sensory modalities, including cross-modal animal experiments. Finally, the majority of synaesthetic reports and accounts include some aspect of vision, usually colour or spatial qualities, as the secondary sense.

1.1.1 Primary Qualities and Common Sensibles

Before beginning a discussion of how aspects of visual and auditory sensation may be similar, it is necessary to examine the types of common qualities that may cross sensory boundaries. When discussing such qualities, two terms are often employed: cross-modal qualities and suprasensory qualities. In this thesis, the term of *cross-modal* will be applied to the matching of those sensory characteristics which are experienced in two distinct modalities, while *suprasensory* will refer to those sensory characteristics which are independent of a single modality. Drawing cross-modal correspondences may certainly involve suprasensory qualities; thus, these distinctions may often overlap. For example, matching between auditory pitch height and visual luminance is cross-modal: the former is exclusively aural, the latter visual. Spatial characteristics of a sensory experience are an example of the suprasensory—they may be experienced through several modalities, independently or simultaneously.

The qualities that are suprasensory have been the topic of much debate. Aristotle (4^{th} Century BC) stated that rest, motion, number, size, unity, and shape comprised a set of characteristics known as *sensus communis*, or *common sensibles*. Locke (1690) further developed this concept of common qualities into those that reflect physical qualities and those that are not perceptions of the actual physical qualities. These he termed *primary* and *secondary qualities*, correspondingly. Locke listed the primary qualities of objects as solidity, size, shape, motion, and number. Marks (1978) points out that these qualities are not only characteristics of the physical world, but also originate in the sensory experiences of the external world. Primary qualities are sensory reflections of an objects' true physical qualities. Secondary qualities, then, are perceptions that are not representative of the true physical nature of the object or environment. One example of this is the different perceptions of a red ball. The tactile and visual perception of the ball are spherical, and thus a quality of its actual physical state. However, the perception of the ball's colour is actually a psychological construct formed from the separate wavelengths reflected from the sphere under certain conditions, absorbed in the retina, and thus not an intrinsic physical quality of the sphere. One of the most important features of this distinction is that primary qualities are available to multiple senses, and are thus suprasensory; secondary qualities are available to one sense only. Galileo (1623) had outlined a smaller list of primary qualities, consisting of size, shape, quantity, and motion.

For the purposes of this thesis, no attempt is made to justify the use of any one particular set of primary qualities. Instead, it is accepted that such suprasensory qualities exist, but the question as to whether such qualities are universal remains. There are two possible explanations for the presence of suprasensory qualities in perception: either the understanding of such qualities is an innate aspect of perception, or suprasensory qualities of perception evolve. The ability to evolve such *categories of understanding*, or *schemata*, to use Kant's terminology (1781), may provide a psychological advantage in development.

It is difficult to provide evidence to firmly support either of these hypotheses. Studies demonstrating that animals experience cross-modal transfer in learning tasks (Marks, 1978) demonstrate that suprasensory representation or cross-modal transfer of sensory information is not purely a human phenomenon. Studies with infants (Bruner & Koslowski, 1972; Meltzoff & Bornton, 1979) provide evidence that these same abilities are to some degree innate. These representations and transfers may serve an active role in the development of a cognitive system. Indeed, amodal perception, or infant synaesthesia, may be a normal stage of development (Maurer, 1993; Baron-Cohen, 1996).

1.1.2 Brightness as a suprasensory quality

Accepting that sensory qualities which cross individual modality boundaries do exist, we may begin to examine which qualities are analogous between senses. These considerations will be primarily limited to those between vision and audition, as they occur most regularly in previous literature. Additionally, a large majority of recorded synaesthetes report visual-auditory correspondences. In the majority of cross-modal literature, *brightness* recurs as the most commonly hypothesised suprasensory attribute outside of those traditionally included in primary quality or common sensible sets. Hornbostel (1931) asserted that brightness was a true suprasensory attribute, present in visual, auditory, and olfactory perception. Other early studies equate visual brightness with auditory pitch (Rich, 1919; Hartshorne, 1934) making it a analogous cross-modal quality. One key reason for focusing on brightness as a suprasensory quality is the frequency with which it is found in synaesthetic correspondences. Colour imagery is reported as the synaesthetic percept most frequently paired with an evoking stimulus. Brightness plays a key role in describing such imagery (Marks, 1975; Cytowic, 1989). The synaesthete is the embodiment of cross-modal matching *in stasis*. If cross-modal mechanisms are at play in developing synaesthesia, common characteristics in various synaesthesiae may provide information as to the nature of these mechanisms.

The precise meaning of *brightness* within cross-modal studies is problematic. As a suprasensory psychological construct, this quality may or may not directly represent a physically measurable property. Due to the common usage of the adjective *brightness* in descriptions of primary visual events (but also auditory events), the same term is used to represent a number of qualities in different studies. In one study it might be used as a synonym for the physically measurable quality of luminance; in others, it may be a purely psychological construct independent of modality.

In order to alleviate potential confusion, it is important to individually examine exactly how the term of *brightness* is employed in each study. In Hornbostel's (1931) experiment, subjects were asked to match auditory and olfactory brightnesses of different events together, as well as matching grey surfaces to odor. Subjects were able to make regular matchings between increasing luminance (grey patches, closer to white) with the same odors that had been matched with increasing auditory pitch height. This approach posits brightness as a purely psychological construct independent of physical properties for hearing and smell, while at the same time asserting that such a construct directly correlates to increasing luminance in vision. This fundamental assertion that increasing brightness corresponds to increasing luminance is at work in most experiments. As a visual quality, other physical characteristics may play a part in perceived brightness.

As a suprasensory quality, it may be said that brightness is closely correlated with luminance for most subjects. It should be noted that most cross-modal experiments examine matching performances with visual stimuli that vary monochromatically (white–black, varying grey levels), using printed images or projected lights. With only this monochromatic information as visual stimuli, other factors that may contribute to brightness are never considered in analyses or discussions. Additionally, if the underlying assumptions as to the nature of visual brightness are flawed, so then are the conclusions about brightnesses of sounds and smells, if such conclusions were drawn from matching these qualities to visual stimuli.

Bearing these caveats in mind, there is a long history of cross-modal matching experiments that examine the role of brightness in multiple sensory modalities. Hornbostel (1931) did manage to demonstrate regular matchings between visual patches of varying grey level with odors, auditory tones with odors, and the same visual patches with the auditory tones. The regular nature of the matchings when analysed together provided evidence for the presence of a quality of *brightness* in olfaction and audition, and that this common quality dictated the matchings. Cohen (1934) tried and failed to replicate Hornbostel's results; but, as Marks (1978, p. 61) points out, that failure may be due to a strictness in analysis that may ignore the underlying nature of the subject's matching performance.

Wicker (1968) performed a colour-tone comparison experiment requiring subjects to judge similarity between coloured paper and musical tones. Subjects also judged similarities of the paper patches to one another, and tones to one another. He then produced a two-dimensional similarity space of tones and colours. Wicker incorporated the use of adjective-scale ratings given to the colour stimuli to aid in the identification of the axes for this two-dimensional space. He suggested two potential axis pairs for a similarity space and a semantic space, as no clear set came from the data: *Density-Vividness* orthogonal to *Volume-Heaviness*, or alternatively *Pitch-Brightness* orthogonal to *Loudness-Contrast*. One of the problems in analysis came from the introduction of varying hue rather than simple patches of varying monochromatic grey levels. If increasing *brightness* is truly equated with increasing luminance, such a hypothesised pitch-brightness association should be evident by examining the cross-modal comparisons using luminance data for the colours, but no such analysis was performed.

Marks and Stevens (1966) did examine subjective brightness in relation to luminance and determined that it follows the psychophysical power function:

$$\psi = k(L - L_o)^{\beta} \tag{1.1}$$

"...where k is a proportionality constant whose value depends on the choice of units of measurement, β is the exponent, and L_O approximates the absolute threshold." These subjective brightness functions were again formulated from experiments using varying levels of white light as visual stimuli. In cross-modal matching experiments, they examined the regular matchings formed between varying luminance and loudness of auditory tones by adjusting either the light or tone to match the other. Their results revealed that cross-modal matching of auditory loudness and visual luminance also follows the power law. Equally important is the result that consistent matchings were found by matching visual brightness and auditory volume. Therefore, regular matchings may be performed using more than just the previously hypothesised equivalence of pitch height and brightness.

Marks (1974) extends this idea in an experiment matching brightnesses of audio tones with white lights. He found that brightness increased with pitch height for tones of constant loudness and increased with loudness for tones of constant frequency. Marks hypothesised that auditory brightness was then composed of both components: high brightness included high pitch and high loudness. Additionally, he suggested that volume was an orthogonal dimension to auditory brightness, with high volume composed of high loudness and low pitch, and low volume composed of low loudness and high pitch. Marks later suggests that:

> *There is even an alternative conceptualization, namely, that pitch and loudness are themselves components of a third, unitary attribute—* density*—and that density is the singular and primary auditory correlate of visual brightness. It is intriguing that the attribute now known as auditory density (see Guirao & Stevens, 1964; Stevens, Guirao & Slawson, 1965) once was called "brightness" (see Boring & Stevens, 1936). (Marks, 1989, p. 587; italics mine)*

This demonstrates that Marks' results have led to a belief in cross-modal correlates, rather than a singular suprasensory representation of brightness.

Marks noted that not all subjects responded in an identical manner when given the matching task, even using monochromatic visual stimuli. Subjects may easily invert their matching criterion when requested and perform matchings in a completely inverse manner. Additionally, some individuals' preference in assigning matches between loudness and visual brightness alters with the contrast between background and foreground for the visual stimuli. No clear theory is provided for why some subjects make opposite assignments between modalities in some cases.

In a later set of experiments (Marks, 1989) subjects were allowed to match auditory stimuli that varied in frequency and loudness with white lights. He found that subjects used different criteria in making their matches, with 50% using frequency exclusively to make their matches, 25% using loudness exclusively, and 25% using a combination of the frequency and loudness (auditory density). This sheds more doubt on the suprasensory nature of visual and auditory brightness.

1.1.3 Innate nature of Cross-Modal brightness transfer

The earlier mentioned studies were undertaken with mature adults, and regular cross-modal matchings were observed. Experiments with infants provide information into the types of cross-modal transfers possible in an immature system. Lewkowicz and Turkewitz (1980) found that infants responded to auditory loudness after becoming familiarized with different lights. This was done by monitoring the infants heart rate when presented with different intensities of sound and light. Those stimuli in a single modality that the infants were most familiarized with produced the least change in heart rate; novel stimuli caused an increase in heart rate. Therefore, the smallest change in heart rate with presentation of stimuli to another sensory modality was said to indicate a cross-modal match between the new stimulus and the stimulus with which the infant was familiarized.

Cardiac measures indicate that infants matched a light of 39 cd/m^2 with a tone of 74 dB by demonstrating increased heart rates for tones louder or softer than 74 dB. Maurer (1993) notes that this is the same selection that most adults will make when requested to chose the most appropriate match using the same stimulus sets. Infants' cross-modal matching changes when familiarized with a brighter light. After training with that light, they demonstrated a preference for the louder tone. This provides some evidence for an underlying absolute nature of cross-modal matchings. Unfortunately, no results were produced by varying the light's luminance and determining the matching frequency.

Maurer (1996) informally reports a cross-modal experiment with children between 30 and 36 months old. They were shown videotaped images of a white or grey ball bouncing in synchrony with a high- or low-frequency tone that sounded when the ball "bounced" at the bottom of the screen. As she predicted, the children matched the white ball with the high-frequency tone, and the darker ball with the lower-frequency tone. The image varied in another dimension, however: the white ball subtended less than half the visual angle of the grey one.

This brings up another quality that is often reported as cross-modal: that of size. While not examined at length in this thesis, the size of synaesthetic percepts is almost as agreed upon by visual synaesthetes as the brightnesses of their percepts. That is, louder sounds evoke larger percepts, or *photisms*, than the darker ones (Karwoski & Odbert, 1938). Additionally, the size of photisms varies with the auditory dimension of pitch frequency (Mark, Hammeal, & Bornstein, 1987). Marks (1975) hypothesises that loudness maps onto two visual dimensions: *brightness* and *size*. If pitch height is mapped onto size in one experiment and onto brightness in another, the case for either of these dimensions actually being a suprasensory representation becomes muddled.

Due to the paucity of studies examining infant matchings using audio stimuli which vary in frequency or auditory density, it is difficult to hypothesise whether there is an initial suprasensory quality at work in these tasks or that the ability to formulate cross-modal matches themselves is innate. However, regular matchings do demonstrate that at least the ability to transfer information from one sensory modality to another is innate.

1.1.3.1 Doctrine of Common Psychophysical Properties

The type of cross-modal matching performed by neonates, as described above, may be best accounted for by Marks' (1978) *Doctrine of Common Psychophysical Properties*. This doctrine attempts to account for cross-modal matching made between sensations with the closest common psychophysical sensory qualities (*e.g.*, *intensity* and *duration*). This contrasts to the earlier mentioned *Doctrine of Equivalent Information* which posits sensory representations that are independent of any one sense, and the *Doctrine of Analogous Attributes* which posits that cross-modal matching occurs by aligning scales from different modalities that are analogous in nature.

It may be that people use the term of brightness to describe sensory qualities when no adequate word to describe the appropriate quality exists. *Brightness* may begin as primarily a psychological representation of intensity, but as time and experience accumulate it includes other additional features of visual sensation, including affective components. The naissance relationship with luminance may continue to be observable in experiments, with a potential for other components to obscure that relationship in some cases. If this is so, that same term may be employed to describe sensory qualities in other modalities. Visually bright experiences are more intense, and one particular person may feel that they are more pleasant and attention demanding. When presented with an auditory experience that is more intense, pleasant and attention-demanding to the perceiver, a search for an adequate adjective may yield the word *brightness* from vision. This might be anticipated if vision dominates the other senses.

All of these caveats rather beg the question of what *auditory brightness* actually is. *Visual brightness* is directly correlated with increasing luminance; however, no one, two, or several auditory dimensions reliably predict auditory brightness ratings of even slightly more complex stimuli than those found in the laboratory. This thesis does not attempt to use auditory brightness as a defined quality. Instead, it assumes that many auditory qualities may be evaluated by using visual brightness. Such a matching does not require the existence of either a single suprasensory *brightness* scale or an auditory brightness scale analogous to a visual brightness correlate.

1.1.4 Can music have different hues?

Researchers have also endeavoured to find cross-modal similarities that may align musical tones with the visual qualities of hue. Again, at least a partial impetus for finding such agreements comes from the reports of synaesthetes who experienced "coloured-audition". Hue figures as a key characteristic in most of their descriptions of synaesthetic photisms second only to brightness. One important caveat is that synaesthetes differ in the actual musical constructs that they associate with colour. Rimsky-Korsakov associated musical keys with different colours (Slonimsky, 1955); Scriabin composed "Prometheus" to use different colours for different single notes in an attempt to communicate his perceptions with an audience (Marks, 1975; Peacock, 1985); Messaiaen associated colours with modes (Bernard, 1986), along with other basic musical constructs. Other reports of lesser-known coloured-hearing synaesthetes also show a variety of musical constructs paired with colour (see Cytowic, 1989, for many of these accounts).

Those musical elements that researchers attempt to align to hue—typically pitch height and meter—may be readily ordered objectively in a linear fashion. Colours are considerably more difficult to order objectively, especially when hue is the only component being used in such an ordering. Simpson, Quinn, and Ausubel (1956) performed an experiment that sought out pre-existing correspondences between hue and music. They asked a group of 995 Illinois primary school students to respond with the colour that they "thought of" after hearing a musical tone. In fact, the students were provided with inch-square strips of coloured paper, so that the students may have matched the presented colours that they were presented with rather than an imagined prototypical colour. The results showed that the students matched *yellow* and *green* with the higher frequency tones, *red* and *orange* with the middle frequency tones, and *blue* and *violet* with the lower frequency tones. These matchings may be explained through brightness matching rather than hue-pitch height equivalencies. The order of colours corresponds with the photopic curve of relative visual sensitivity, with psychologically brighter colours matched with the higher frequency tones. Thus, the results may provide no new information about hue's importance in cross-modal matchings as a separate sensory feature.

Polzella and Kuna (1981) asked college students to match excerpts from G. F. Handel's 12 *Concerti Grossi* and match them with *red, yellow, green,* or *blue* without presenting any physical representation of the colours. They reported that the pieces of major tonality evoked primarily *yellow* and secondarily *green* matchings; minor tonality pieces evoked primarily *blue* matchings. Polzella and Kuna refer to these imagined colours as "photisms", implicitly making the rather large and contentious leap of suggesting that the subjects manufactured synaesthetic imagery during the task. They assert that the different key areas evoke photisms of a particular hue, while subjects may have been using another characteristic of the imagined colour for matching. Brightness again figures as a likely characteristic: major keys may evoke brighter imagery than minor keys, and an imagined *yellow* or *green* is prototypically brighter than *blue*. If brightness is the actual quality governing the matches, this result may say more about the brightness ratings of tonality rather than any specific hue evoked. Since synaesthetes often report the brightness components of their percepts before other factors, and also demonstrate quite diverse idiosyncratic hue associations. The potential for the use of visual brightness in a tonality matching task will be explored in Chapters 3, 4, and 5.

Polzella and Biers (1987) extended these findings by obtaining responses from over 100 subjects, matching the same colour set with 12 of the 24 *Preludes* from J. S. Bach's *Well-tempered Clavier*. The pieces were selected for tempo and meter, using equal numbers of quadruple and triple meters, and slow, moderate, and fast tempi. Their results reveal higher proportions of *yellow* responses for fast tempo, *red* for moderate tempo, and *blue* for slow tempo. Again, the authors refer to these responses as "musically induced chromaesthesia" and assert that coloured photisms can be "reliably evoked by recorded musical compositions." The brightness quality of imagined colours might explain such findings if subjects are again aligning prototypically brighter colours with faster tempi.

1.1.5 Common spatial nature of visual and auditory sensations

In addition to brightness, other aspects of visual and auditory sensation have been suggested as suprasensory. One of the suggested qualities that recurs is that of spatial location. Both vision and audition both have a strong spatial component in their separate experiences. The spatial adjectives of "High" and "Low", used to describe musical tones in the most fundamental terms, provide some evidence of the close relation between the sensory dimensions. Physically, there is no apparent connection between the frequency of a sound wave and its location in space.

Bernstein and Edelstein (1971) postulated a common psychological space in which both auditory and visual experiences are represented. This was based on reaction time experiments showing redundancy gain when visual stimuli were presented with ipsilateral auditory cues and redundancy loss for contralateral auditory cues; a cross-modal variant on the classic Posner attention experiment.

Auerbach & Sperling (1974) further tested this notion of a common space hypothesis by asking subjects to determine whether two consecutive stimulus presentations were in the same or different spatial locations. The stimulus pairs could be inter-modal (audio and audio; visual and visual) or cross-modal (audio followed by visual; visual followed by audio). Subject performance led the researchers to conclude that auditory and visual spaces were represented in a single psychological construct. They went further to say that directional perception is thus independent of modality. This predicts that experiences from the other untested sensory modalities should share the same representational space, although further experiments have not been performed with olfaction, proprioception, or other senses. A common space model eliminates the need to account for the process of drawing correspondences between two disjunct spaces in adult perception; however, the question of whether spaces begin disjunct and integrate into a central representation in the process of development still remains.

Another indicator of a potential suprasensory representation comes, as it did with brightness, from the reports of synaesthetes. The size of photisms is second to brightness as the most agreed upon synaesthetic percept quality (Marks, 1975; Cytowic, 1989). Size also has some properties which makes it spatial in nature. Reports from synaesthetes also include detailed descriptions of the dynamic movement of photisms (Cytowic, 1989; *Horizon*, 1995). A notable example of the importance of spatial qualities in photisms is recorded synaesthete DS. When listening to music that she especially likes, the photisms alter from moving in two dimensions (vertical and horizontal) to three dimensions, moving towards and away from her. In this case, the proximity dimension becomes integrated with affect. (1994, *personal communication*).

Spatial qualities of other synaesthesiae, including visual photisms from pain (Steen, 1995, *personal communication*; Cytowic, 1989) as well as categorical stimuli (Galton, 1907; Cytowic, 1989; also, see Section 2.7), and non-visual synaesthesia, such as a geometric-gustatory synaesthete (Cytowic & Wood, 1982b). The commonality of spatial components in synaesthetic percepts and the regular matchings made by normal subjects suggests that some spatial qualities may have at least analogous representations in multiple modalities.

1.1.6 Search for accurate "visual metaphors" for sound

Walker (1978, 1981) investigated the *visual metaphors* that people use to represent musical characteristics. Written musical notation supplies one set of rules for visually representing musical event. Walker allowed subjects to choose which visual properties best comprised a personal notation for music. He argues that the visual height association with pitch height is a function of musical training (Walker, 1987). However, these conclusions are based upon subject choices when supplied with a set of two-dimensional representations, rather than the more ecologically valid spatial task discussed in Section 1.1.5. This finding does call the effect of training and experience with cross-modal associations into question.

Walker's study (1987) sought to examine this effect by obtaining *visual metaphor* matching with musical characteristics with musically trained and untrained Canadians, and compare those responses with Innuit, Haida, Shuswap, and Tsimsian subjects. Musically trained subjects responded most consistently with their choices of *visual metaphors*. The apparent reason was that they already had a representational written notation, and compared other notational systems to their fixed one. *Frequency* was matched with vertical placement of symbol, and *duration* with the horizontal length of the symbol. Two qualities not normally represented with symbols, *amplitude* and *waveform*, were matched with symbol size and complexity of visual pattern, correspondingly. Walker's attempts to explain the irregular differences between the cultural groups are quite complex, and reveal little about how cultural differences might change cross-modal associations.

1.2 Integral and interacting dimensions

Several researchers have examined the interaction of sensory dimensions rather than seeking common characteristics between them. The *Garner Interference Paradigm* (Garner, 1974) was used to examine such interactions. This paradigm states that when presented simultaneously with two *physical dimensions*, subjects attempting to classify attributes of one dimension will experience an intrusion from the other dimension if the two dimensions interact, similar to *Stroop Interference* (Stroop, 1935). The level of this intrusion, or *Garner interference*, may be measured as the difference between a baseline measure of classification speed for the single dimension and speed when presented with an accompanying independent dimension. Pomerantz and Garner (1973) tested the intrusion of parenthesis pairs, where subjects were required to classify only one of the parentheses while ignoring the second. This design allowed for stimuli composed of *positively correlated dimensions* [(( ] and [ ))] as well *negatively correlated dimensions* [ ) ( ] and [( )]. Performance with negatively correlated dimensional stimuli can reveal Garner interference with the classification task, while performance with positively correlated dimensional stimuli can reveal a redundancy gain in classification speeds.

Melara and O'Brien (1987) used a similar approach to examine possible Garner interference between pitch height and vertical position of a visual stimulus. This poly-modal Garner interference they termed as *Cross-modal Interaction*. They did find Garner interference when the second dimension was uninformative to the classification task. However, they found redundancy gain for the positively correlated dimensions of pitch height and visual height, but observed no redundancy loss for the negatively correlated dimensions. Melara (1989a) concludes that the observed interaction may represent yet another form of "dimensional interaction"—in this case, a cross-modal interaction. However, if these physical dimensions have suprasensory representation, the modality of presentation should not change the interaction to such a degree as to eliminate the intrusion from negatively correlated dimensions.

Melara (1989a) continued to examine this cross-modal interaction with a similar task using the physical dimensions of pitch height and colour. Colour here, again, was defined in monochromatic terms: either *white* or *black* circles were used as visual stimuli. Again Garner interference effect was observed, as well as redundancy gain with no accompanying redundancy loss. These results suggest that neither vertical position nor colour are represented in the same manner as pitch height, but that these representations do interact cognitively.

Melara & Marks (1990c) continued to examine the interactions of perceptual dimensions, using Garner interference as a technique to establish whether dimensions were *integral*, *separable*, or *corresponding*. *Integral* dimensions are those which are inseparable in perception. Brightness and saturation of colour are suggested as one possible set of integral dimensions (Garner & Felfoldy, 1970). *Separable* dimensions, then, refer to perceptual dimensions that do not interact. Interactions of dimensions from more than one modality, such as the pitch height–visual position interaction (Melara, 1989a) were designated as *corresponding* dimensions. It follows that measures of Garner interference should provide evidence for whether a particular two-dimensional stimulus set consists of integral or corresponding dimensions. Strong Garner interference indicates integral or corresponding dimensions; no Garner interference indicates completely separable dimensions. Corresponding dimensional interaction does not make any claims upon whether such sensory attributes are common (*i.e.*, suprasensory): only that the physically measured dimensions interact.

1.2.1 Locus of cross-modal transfer

If such dimensions do interact cognitively, at what stage of processing do they interact? If they are suprasensory, they interact at a holistic level. If they are cognitively similar representations, but not identical in nature, they interact at an analytical level. Melara and Marks challenge the view that perceivers have no access to the dimensionality of stimuli in the initial stage of perception, and that such dimensionality is discriminated only with cognitive effort (Lockhead, 1972, 1979). Rather, they believe that access to the *primary* dimensions of stimuli is immediately available to the perceiver. These primary dimensions are those that all perceivers align with the same axis orientation. This level of processing they term as *attribute-level processing*. They hypothesise that Garner interference occurs at the next level up, termed *stimulus-level processing*. Such a process is the actual formation of a perceptual whole from the primary dimensions. Thus, cross-modal interaction would follow the *Doctrine of Analogous Attributes*, rather than a model using Aristotelian common sensibles.

1.2.2 Is language necessary for cross-modal transfer?

In Melara and Mark's model, language is not necessary for stimulus-level processing, but might play a part in the formulation of the dimensions that subsequently interact. These primary dimensional scales must either be innate or evolve. Additional studies (Melara & Marks, 1990a, 1990b) examined cross-modal interactions with language components as part of the stimuli set, and cross-modal interaction is observed with these language elements. However, cross-modal transfer experiments performed with neonates (Lewkowicz & Turkewitz, 1980; Maurer 1993) as well as animals demonstrate that language is not a requirement for transfer. The types of transfer available without language may be of a particular nature: the matching of common psychophysical properties. Higher-level types of cross-modal transfer may subsequently be possible only with language skills. Certainly, metaphoric description requires language.

Lewkowicz & Turkewitz (1980) used heart rate monitoring techniques to demonstrate intensity transference between light and noise for one-month olds. Most interestingly, the most effective transfer occurred at the same level that adults "matched" between the two modalities (74 dB white noise with 39 cd/m^2 white light), providing evidence for the innate nature of some dimensional representations as well as the mechanism for drawing cross-modal correspondences. Meltzoff & Borton (1979) found that one-month olds visually attend to pictures of pacifiers that they have explored orally. Walker-Andrews & Gibson (1986) found the inverse relationship: the same age infants attended to the pacifier picture that they had *not* explored. Maurer (1993) was unable to replicate this effect for one-month olds, but points out that although cross-modal transfer may be innate, it may also be ephemeral and obscured by the infant seeking stimulation across all sensory modalities. She also points out that there are optimal ranges of stimuli intensity for the tasks, and that these change as the testing environment changes. Maurer explains the transfer here in terms of *continuous* and *discontinuous* energy, accounting for transfer in a perceptual system that is hypothesised to be undifferentiated. Therefore, the rough "energy" pattern from each stimulus is compared without an awareness of its source modality, following the *Doctrine of Common Psychophysical Properties* (Marks, 1978).

Baron-Cohen (1996) draws a distinction between two types of hypothesised neonatal sensory experiences underlying the formation of synaesthesia. These are the *Cross-Modal Transfer* (CMT) and *Neonatal Synaesthesia* (NS) hypotheses. The CMT hypothesis accounts for the transfer between modalities in infants as a result of representing sensory experiences in an abstract form, regardless of the experience source modality. When another experience is represented with a similar abstract form, cross-modal agreement is made and cross-modal transfer can take place. The NS hypothesis states that until approximately four months of age, sensory experiences are undifferentiated. Input received from a sensory organ is experienced tactually, visually, and auditorily. Cross-modal matching subsequently occurs when two experiences evoke similar neural activity. Thus, the *Doctrine of Common Psychophysical Properties* may have a direct physical counterpart in *The Doctrine of Neural Correspondences* (Marks, 1978).

The CMT hypothesis may provide a basis for understanding cross-modal transfer in mature individual, while the NS hypothesis may account for synaesthetic development. Even if the NS hypothesis is true, we still must account for such on-going transfer. These two hypotheses, and their possible role in the formation of synaesthesia, will be more fully discussed in the next chapter.

One might assume the necessity of language in formulating cross-modal correspondences from the terminology used to describe such correspondences in cross-modal literature. In particular, the term *semantic* is used in an irregular fashion. Melara & Marks (1990a) use the term *semantic crosstalk* to describe a level of interacting dimensionality in Garner interference separate from *perceptual crosstalk*, which occurs at a lower level of processing. Such higher-level interaction may require language to formulate the concepts that interact. *Semantic* is used in these studies to refer to deep meaning, rather than any quality specifically language-based.

Even so, language may still act as one (or several) of the dimensions in cross-modal transfer and interaction, and may provide important clues as to the nature of those dimensions. Melara (1989) points out that the labels of "HI" and "LO" for visual position *and* pitch frequency does provide clues as to the common nature of their spatial scales. However, Pomerantz (1989a) did not find Garner interference with pitch height and the presented words "HI" and "LO". Such an interaction may occur at a higher cognitive level, and thus is less pronounced. The observed Garner interference between visual height and pitch height may be due to the hypothesised common visuo-auditory space, accounting for the development of the same written labels in many diverse cultures.

Marks (1978) points out that while "bright" was initially a term of luminosity, it was being used to describe sounds as early as 1000 AD. The transference over time of descriptive terms to other modalities provides clues into an ordering of senses. Marks (1978) also notes that the etymology of the word "sharp", recorded in the *Oxford English Dictionary*, provides another such clue. First used as a tactile adjective, it subsequently was used as a visual adjective, then as an auditory adjective. Williams (1976), in the course of linguistic research, determined that adjectives from one sensory modality could transfer aspects of their meaning to other modalities, but only in a limited fashion. These transfers were primarily unidirectional, with the exception of *colour* and *sound*, which can use each others adjectives. He believed touch to be the most primary of senses, with the terms used to describe touch readily transferred to other sensory modalities. In the same fashion, the primitive senses can pass their terminology onto more complex ones, but that the transference is unidirectional. The ordering is shown schematically in Figure 1.1.

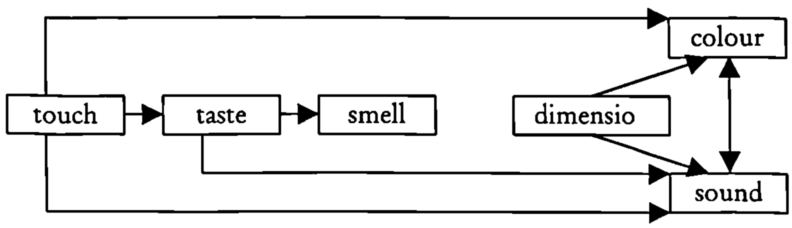


Figure 1.1. Adjective transfer between sensory modalities. (Williams, 1976).

While these are not hard and fast rules (poetry sometimes does not follow these rules), it does indicate how these adjectives transfer for most of English usage. This schematic may provide useful insight into how easily non-language information may transfer between sensory modalities. The importance of sound and colour as the senses that receive the most transferred information follows from their presence in most synaesthesiae.

Köhler (1947) performed an experiment that examined a simple cross-modal matching between nonsense words and simple line drawings. The words *Maluma* and *Takete* were used, and matched with line drawings similar to those found in Figure 1.2. *Maluma* was matched with the rounded drawing on the left; *Takete* with the angular drawing on the right. Although an explanation of exactly which processes occurred in the subjects was not attempted, a strong case is made for the possibility of simple language constructs to transfer to other sensory representations. As the words were presented orally, it is unlikely that subjects used the written form of the word as their matching criterion—unless, of course, subjects visualised the word spellings. Nevertheless, this simple experiment demonstrates that some low-level aspects of language may transfer or be represented across sensory boundaries.

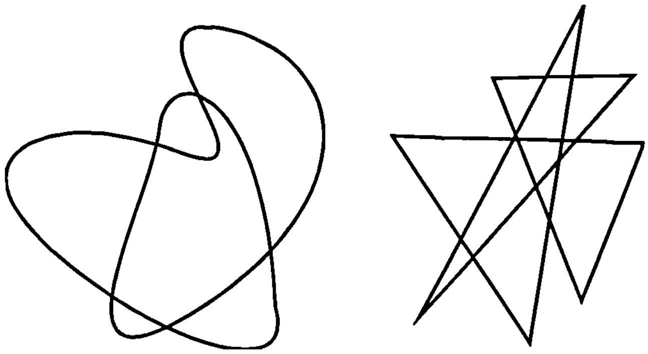


Figure 1.2. Graphic shapes similar to those for matching with the words "Maluma" and "Takete" in the K¬hler experiment (1947).

The next chapter will discuss the role of language constructs (letters and phonemes) in the development of several synaesthesiae. The role that language plays in a large proportion of these synaesthesiae suggests that it may help develop these most fixed of cross-modal correspondences.

1.3 The approach of the thesis

This thesis will take the position that the process of drawing correspondences between sensory modalities is a dynamic, continuing process in the developing individual and mature adult alike. This process is at work whenever the individual is presented with novel information, and seeks to utilise it in some manner that requires discriminating between dimensional levels of that information. Such a process may be fundamental to the formulation of schemata. It also helps to explain individuals' willingness to readily use cross-modal information with little guidance. For example, children willingly draw shapes in response to sounds without suggestions of what qualities of sound they are trying to represent (Chacksfield et al, 1973). Finally, it gives some added insight into the importance of metaphor in language.

1.3.1 The SoC hypothesis of cross-modal perception

During perception, the sensory system must initially make a measurement of a stimulus. This measurement bears a monotonic relationship to the stimulus being measured. One example of this type of measurement is the magnitude of a stimulus. A measurement on its own has little meaning; however, measurements may be compared to others within their own domain. This comparison gives more meaning to the measurement, due to its comparison with an expected stable population of measurements. Any measurement of a stimulus is thus judged as being typical or atypical to the population norm. That is, a measurement may be a likely or unlikely event. This comparison of the sensory measurement to a stable population performs an *evaluation* of the measurement.

Andrews (1964) originally conceived of this process as an error correcting mechanism for adjusting biases that a sensory system is subjected to. Watt (1988) extended this idea by demonstrating how such a mechanism operates more generally in the context of the visual system. This idea explains why a person with degrading vision still sees sharp edges to objects rather than blurry edges. The edge detected is compared to the recent population of all encountered edges. If the object is large and distinct enough, the person will report seeing a sharp edge.

The brightness of a perceived light provides another example of this mechanism in simple terms. The initial measurement of the light produces some value of the stimulus, such as a magnitude of *8.7*. This measurement is subject to systematic errors, and has no meaning by itself until it is compared to the stable population of other magnitudes of lights which have been encountered. This measured light magnitude may be two standard deviations greater than the mean magnitude, and thus is *evaluated* as being an unlikely event: the light is very bright to the perceiver. The population of encountered magnitudes changes with time, so after repeated presentations, the initially bright light becomes the average intensity level.

But what occurs when the individual does not have a stable population of measurements to compare a sensory measurement with? When initially presented with a new set of sensory information, no scale exists with which a perceiver can make an evaluation. In such a case, the perceiver may use another existing scale to make the evaluation. They "borrow" a scale from another schemata to compare the measurement with—a *Scale of Comparison* (SoC). For example, a person attempting to understand varying pitch height of musical tones early in development may hear a note and ask the question, "If this were a light, would it be a bright light or a dim light?" That is, the person attempts to understand the measurement in relation to those scales she already knows about.

The initial evaluation of new information only needs to be approximate: if the measurement were a light, would it be *very* bright, *very* dim, or of an average intensity. Over time and repeated encounters, the evaluations of information become more refined. The individual becomes more accurate in mapping the judgments onto the SoC for evaluation. Through this process, a separate scale for the evaluation of information develops. The new scale may contain qualities of the stimuli that cannot be represented adequately using the SoC.

This model has some strong implications. Firstly, using SoCs is an on-going dynamic process in the developing and mature individual. It aids in the creation and refinement of evaluative schema in terms already understood by the individual. As scales develop from other scales, a hierarchy of scale inter-relationships develops. The distance and inter-relatedness of two scales from one another in this hierarchy will be referred to as their *cognitive distance*. Scales that are proximal (having small cognitive distance) will share common perceptual and cognitive qualities; distal scales will have fewer of these qualities in common.

Secondly, individuals should most often use SoCs from the same modality as the newly encountered information, but this limitation is not required. A scale will be used as a SoC if it shares some quality, ranging from psychophysical to affective, with the new percept. Scales evolved from the same sensory organ will most likely have the most in common with new information to that same modality. Repeated use of SoCs produced from the same sensory organ may then increase modularity between senses.

Thirdly, the use of cross-modal experiments sets up scenarios which limit the subject's choices of SoCs, while allowing the underlying processes to operate as usual. Regular cross-modal matching naturally occurs, since the experimenter has set up the rules that the subject can easily follow. However, nothing in the consistent cross-modal performances demonstrates that the cross-modal correspondences are suprasensory using this model. In order to make a cross-modal evaluation, the subject makes a measurement of the presented stimulus in one modality and compares it with the population of measurements in the other stimulus modality. This evaluation can then be used to perform the cross-modal task, whether it is an adjustment of a stimulus to a matching level or selecting an appropriate match from a presented set of stimuli. The possible SoCs that a subject can use for cross-modal evaluation is limited by experimenter's choice of stimuli.

In the course of a cross-modal experiment, subjects will naturally attempt to align the two stimulus dimensions in a consistent manner. Responses in experiments should not be random, nor will dimensional alignments change in the course of an experiment. Also, the *cognitive distance* between scales—how related the two scales are in their development or sharing of common sensory qualities—can aid in predicting which way a person will align the SoC with the stimulus scale.

The last implication of this model is that repeated use of a SoC can produce synaesthesia if the individual persists in using a SoC for evaluating sensory stimuli. This implication will be discussed in more depth in Chapter Two.

1.4 Summary

The human perceptual system needs to understand the external world in the face of constant sensory input from multiple modalities. In the process of development, transferring knowledge from one modality to another provides an advantage in learning, as well as determining which sensations occur regularly in two or more modalities.

Cross-modal experiments provide some information into which types of qualities are matched across modalities, and the mechanisms that work to formulate these agreements. Despite the significant findings observed, no general theory is agreed upon to explain how cross-modal agreements are created, or whether the agreed upon qualities are suprasensory or merely analogous.

The sensory qualities observed as analogous in a majority of cross-modal experiments were *brightness* and *spatial qualities*. Despite Hornbostel's assertion that brightness was a true suprasensory quality, the difficulties in defining auditory and olfactory brightness, along with the large number of different types of sensory qualities that may be successfully matched with visual brightness suggest that visual brightness may have a number of correlates in other modalities. Findings of subject concordance for hues matched with auditory events were reconsidered, and it is suggested that matching with visual brightness alone may suffice to explain the matching patterns found. Studies into the shared nature of auditory and visual spaces demonstrate that these spaces may be one and the same. Other spatial characteristics, such as percept size, may follow the brightness model, having analogous qualities in other sensory modalities.

The concepts of integral and interacting dimensions were introduced, as well as the potential interference that might be generated by conflicting information in analogous representations. Neonatal and animal experiments provide evidence that cross-modal matching abilities are to some degree innate, and do not necessarily require language abilities to function.

Finally, a model describing an on-going dynamic process for formulating cross-modal agreements was presented. This model posits that an existing representational *scale of comparison* (SoC) may be "borrowed" to aid in the evaluation of a novel stimulus set. Over time, SoCs aid in the generation of fully developed scales which may themselves be subsequently borrowed and used to make further scales. This process of developing new scales creates a hierarchy of interrelated scales. The *cognitive distance* between scales within this hierarchy provides a variable to predict how two scales might be aligned by an individual.

This model operates at an early level, with developing individuals using the mechanism to create new scales from innate evaluative scales, (*e.g.*, stimulus magnitude estimations) that can be independent of modality. It also accounts for more complex cases, such as adults learning to discriminate between musical qualities (*e.g.*, dissonance and chord type) or different types of instrumental timbres. The following chapter extends the SoC hypothesis from normal perception to synaesthetic perception, and describes how an evaluative cross-modal mechanism aids in the development of synaesthetic responses and imagery.

1 THEORIES OF CROSS-MODAL PERCEPTION AND SYNAESTHESIA..1—1

2 Literature Review Of Synaesthesia

2.1 Introduction

In the last chapter, cross-modal correspondences formed by normal individuals were discussed. A model of a dynamic process which continually draws cross-modal correspondences when presented with novel stimuli was introduced. This chapter will review past research into a condition that demonstrates cross-modal correspondences *in stasis*: synaesthesia. The SoC hypothesis will be further discussed with its implications for the development of synaesthesia in children, and its fixedness later in life.

The term *synaesthesia* is derived from the Greek *syn* (συν), meaning union, and *aisthesis* (αισθησις), meaning sensation. The condition is one of literally *joined* sensations. Sensory input to one modality gives rise to sensory experiences in one or several additional modalities. The perceiver regards the secondary percept as a sensation different to one that would be produced by direct sensory stimulation in that modality. That is, synaesthetic visual percepts are unlike actual vision for the synaesthete. One example of a synaesthetic association is a synaesthete who visualises green liquid-like sparkling images and diaphanous coloured clouds when listening to music (Whipple, 1900). Another synaesthete has the sensation of feeling geometric objects when tasting and smelling (Cytowic & Wood, 1982b). Although reports of individuals with synaesthesia date back 300 years (Locke, 1690; Marks, 1975), its low incidence and the relatively small number of formalised research papers on the condition have not produced an agreed-upon definition for precisely what constitutes synaesthesia. The fluctuating criteria for diagnosing synaesthesia have confused estimates of its incidence.

Exacerbating the confusion over what types of joined sensations constitute synaesthesia is an eagerness by researchers to detect synaesthesia where it may not actually exist. Synaesthesia was a condition of fashion early in this century, both in the artistic and scientific communities, producing a variety of different types of studies. Often the reports of synaesthetes anecdotal in nature, and are often single case studies into an individual synaesthete's associations (Cytowic, 1982b; Riggs & Karwoski, 1934; Hart, 1909; Whipple, 1900). Many researchers have termed any associations observed during cross-modal tasks as experimentally "induced synaesthesia" (Simpson, Quinn, & Ausubel, 1956; Polzella & Biers, 1987). Others have attempting to artificially induce synaesthesia with drugs (Simpson & McKellar, 1955). The inclusion of these experiences with synaesthesia case-studies further confuses exactly what is defined as *synaesthetic*.

Terminology used to describe the different forms of synaesthesia is for the most part universal: the name of the secondary sense is given first (sometimes referred to as *imagery sense*), followed by the name of the primary sense which evokes the associated imagery. For example, *coloured-hearing* refers to coloured images that are evoked by auditory input. The high proportion of synaesthesiae that have vision as the secondary sense yielded a term used to describe the synaesthetic visual images: *photisms*. The terms of *chromaesthesia* and *pseudochromaesthesia* have been employed to describe coloured sensations evoked by alphabetical characters and numeric digits, although *chromaesthesia* has also been used to describe any coloured imagery invoked by a primary sensory modality.

In the previous chapter, a model of using scales of comparison (SoCs) for sensory evaluations was used to help explain how normal subjects performed in cross-modal experiments. How might this idea be used to explain synaesthesia, where cross-modal perception continually occurs? The SoC hypothesis states that mature individuals will use an existing scale upon which they map new information to make sensory evaluations. Developing individuals have much more opportunity to use SoCs, as they are constantly encountering new information and developing schemata. Synaesthesia may develop when an individual persists in using a SoC in evaluations; a normal individual would stop using the borrowed scale when a new scale to represent the information had fully developed.

2.2 A definition of synaesthesia

Synaesthesia is defined in this thesis to be imagery that is evoked as a fixed response to a specific stimulus. The stimuli may be either sensory input or what will be called *categorical stimuli*: stimuli that belong to the same conceptual group. Some examples of categorical stimuli are musical sounds, the days of the week, months of the year, numerals, the alphabet, TV stations, school grades, etc. (Cytowic, 1989, p. 201). The evoked imagery can take place in any sensory modality. The evoked imagery is involuntary, consistent, and durable. That is:

1. The associated imagery are involuntarily evoked whether by direct input to a sensory organ or by internal processes that use the categorical stimuli;
2. The same associated imagery is consistently experienced in response to the evoking stimuli;
3. The vividness of the associated imagery does not decay with repeated presentations.

These criteria are a subset of Cytowic's diagnostic criteria for synaesthesia (1989, see Section 2.6). It is important to note that in this definition a synaesthetic association need not be elicited only by an external sensory input, but can also be internally generated.

2.3 Two different forms of synaesthesia

The wide criteria applied to synaesthesia may have led to a state where more than one condition has been studied. The different types of joined sensations in the reports can be divided into two categories. The first of these categories are joined sensations where an entire sensory modality crosses with another. The perceptual modalities involved are directly tied to sensory organs, (such as taste, touch, smell, sight, and sound). These forms of synaesthesia are rare, and may be bi-directional: that is, one sensory input produces imagery in another sensory modality and vice versa. This type of synaesthesia will be referred to hereafter as *diffuse synaesthesia*. Examples of diffuse synaesthesia are MW (Cytowic, 1982b), who experiences all tastes as geometrically felt shapes, JR (Baron-Cohen, 1996), who experiences both sounds paired with vision and visual images paired with sounds, DS (Cytowic, 1989; see also Chapter Seven), who experiences images in response to all sounds, and S (Luria, 1968), who reported multi-directional sensory crossings in four modalities.

The other form of synaesthesia consists of imagery involuntarily evoked from a *selective* subset of input to a sensory modality. That is, a conceptual group of stimuli, or *categorical stimuli*, evoke imagery for the perceiver. Examples of this are visual imagery evoked for only musical tones (and not other auditory stimuli), or alphabetical characters and numerals (and not all written symbols). This type of synaesthesia will be referred to hereafter as *selective synaesthesia*, and is the most common of reported synaesthesiae. Often with selective synaesthesia, the synaesthetic percept is reported to be evoked by the internal use of a cognitive schema rather than direct input to a sensory organ. For example, visual imagery for a synaesthete who experiences coloured letters or numbers is evoked when the input is presented aurally and visually. The synaesthete also experiences the same imagery when merely "thinking of a number" (Cytowic, 1989; see also Chapter Seven).

Some reported associations experienced by synaesthetes, such as synaesthetic forms and personality traits evoked by categorical stimuli are very different in nature from diffuse synaesthesia. *Synaesthetic forms* are spatial imagery experienced by synaesthetes when thinking of categorical stimuli, such as numerals, days of the week, and seasons (Galton, 1907; Cytowic, 1989; see also Chapter Seven). *Associated personality traits* refers to a synaesthete's experiencing a sense of personality ascribed to certain stimuli. For example, some synaesthetes report that a particular number has a gender and a personality of its own, and these remain consistent and constant (Cytowic, 1989). Both these forms of synaesthetic association are further discussed later in the chapter.

2.4 A theory for the development of synaesthesia

The SoC hypothesis for cross-modal perception described in Chapter One posits that novel stimuli can be judged as being likely or unlikely events by comparing a stimulus measurement to a stable population of measurements in another domain. This process of comparing the measurement performs an evaluation of the stimulus. At first, the mapping onto the SoC will be rough, but increase in accuracy with experience. The development of synaesthesia is hypothesised to be a result of continuing to use a borrowed scale of comparison for evaluating sensory stimuli after a new scale for evaluating the stimuli has evolved. It is now necessary to see how the hypothesis can account for both selective and diffuse synaesthesia.

There are two general characteristics of selective synaesthesia. Firstly, certain stimuli exist that can generate multiple and continuous imagery for the synaesthete; that is, imagery can be continuously evoked in more than one modality. Secondly, within one input modality, not all stimuli will have this effect. The SoC hypothesis can explain both of these aspects of selective synaesthesia. If a SoC was used to aid in evaluative judgments for stimuli of a certain conceptual group, continued use of the SoC is hypothesised to eventually produce synaesthetic imagery for that conceptual group. If *all* stimuli to a sensory modality are evaluated on cross-modal SoCs, diffuse synaesthesia results; if only some stimuli are evaluated on cross-modal SoCs, the result is selective synaesthesia.

For example, some synaesthetes only have associations for sounds that are musical notes (see Chapter Seven). This develops because an individual presented with sounds of a certain conceptual group, such as musical notes, will use a SoC for evaluating the notes. That individual might judge the musical notes by their pitch height, and map that measurement onto an existing visual SoC. When making judgments about other types of sound, the individual may use a different SoC. The resulting types of synaesthesia that develop are different due to their differing origins.

Why then do all people not end up as synaesthetes? After a period of time, a scale to evaluate the sensory stimuli evolves in its own right. The synaesthete persists in using the borrowed SoC as well as the new scale for evaluations, while a normal individual would abandon the SoC that was initially borrowed for evaluations. Some individuals may be more predisposed to forming lasting connections when using SoCs.

This hypothesis cannot account for all forms of synaesthesia, however. Up to this point, the stimuli that have been considered are those that vary in measurement along some continuum, such as magnitude. These stimuli can be placed in an order along this continuum. One type of selective synaesthete experiences imagery evoked by categorical stimuli. Such stimuli consists of discrete elements which do not themselves comprise a continuum. Two examples of this are letters and numerals, which may be discriminated but have no independent ordering outside of that supplied by learning the concepts that they represent.

When confronted with this type of stimulus, an individual may assign the stimuli to discrete levels of another developed scale, such as colour. Why should a child use associated colours rather than the numeral, if they can discriminate between different numerals? One reason is that the child may be more accustomed to working with colours, and find operations such as ordering the numerals easier by ordering the colours: the child may find it easier to perform mental operations with the colours.

Synaesthesia may develop during a critical period. If individuals develop scales for evaluating as they encounter new conceptual groups of stimuli, then it follows that the more immature individuals need to develop more evaluative scales. Thus, when the sensory systems are immature, the facility for individuals to use inter-sensory SoCs may be greater. Additionally, individuals with immature sensory systems will encounter more new rudimentary information that they need to learn about. An early developmental period for synaesthesia is supported by the reports that almost all synaesthetes report experiencing their joined sensations as far back as they can remember (Baron-Cohen et al, 1993; Cytowic, 1989; Luria, 1968; see also Section 2.5.1).

With this model in mind, the previous research into synaesthesia will be reviewed, along with the implications that these studies hold.

2.5 Common characteristics of synaesthesia

Many common characteristics of synaesthesia have been reported, although many of these characteristics come from anecdotal reports. The characteristics presented here are those that fit into the definition of synaesthesia outlined in Section 2.3.

2.5.1 Synaesthesia manifests itself early in life, and fixed for a lifetime

Synaesthetes report that their associations have been with them for as long as they can remember (Maurer, 1993; Baron-Cohen, 1996; Luria, 1968; Cytowic, 1989, 1995; Raines, 1909). Early development of synaesthesia is anticipated if it develops as a result of using a borrowed SoC when learning about stimuli from a new conceptual group. Although many synaesthetes do not or cannot account for an underlying association that may have created a synaesthetic correspondence, some can point to specific associations (*Horizon*, 1995). While one synaesthete may have no idea why a French horn has a "green" timbre, another may remember that a particular note is "blue" because the fingering chart she used as a girl had that note in blue (1995, *personal communication*). The report of Vladamir Nabokov telling his mother that the colours of letters on his blocks were "wrong" (Cytowic, 1989, p. 25) has often been used to dispute such a direct correspondence. However, disagreements between synaesthetic percepts and suggested sources for correspondences provide no evidence that such a correspondences were not formed earlier with another source, such as originating from a different primer or toy. Chapter Seven will discuss a four year old boy that demonstrated coloured associations in response to numbers. For this subject, the specific source of those colour associations was located and identified (see Section 7.6.1.1).

Synaesthetic correspondences have been repeatedly recorded as fixed for the life of the synaesthete. The fixedness of synaesthetic imagery results from persisting to use a cross-modal SoC for evaluation or consistently assigning discrete levels of categorical stimuli to levels of an SoC during development. Varying imagery in response to a single stimulus would indicate an unlawful mapping of stimulus onto an associated scale. Consistency in responses makes it possible to use a synaesthete's consistent reporting of associated imagery as evidence for synaesthesia itself (see Section 2.7.1.2). One experimenter presented an experiment wherein a 17-year-old coloured-hearing boy was given the task of learning to associate colours with musical notes (Rizzo & Esslinger, 1988). The colours chosen for assignment differed from the ones that the subject already associated with the stimulus notes. This task resulted in the boy incorporating the new colour into his synaesthetic percepts, rather than having the colour supplant his existing association. One example of the resulting percepts was his report of "Bright red with a strip of yellow" when asked to associate *yellow* with the note *C*. *Bright red* was his previously reported colour association for the note. He was also able to remember these associations after only one presentation of each paired stimulus, further demonstrating the mnemonic power of some synaesthesiae. The synaesthetic associations of Luria's mnemonist S. were repeatedly recorded as consistent over periods in excess of 15 years (Luria, 1968).

2.5.2 Synaesthetic percepts are often selective within the input modality

Synaesthetic correspondences do not always pervade entire sensory modalities; in selective synaesthesia only a subset of experiences in one sense will elicit synaesthetic percepts in the associated sense. Thus, not all sounds will evoke synaesthetic percepts for some synaesthetes. With diffuse synaesthetes, such as the mnemonist S. (Luria, 1968) or geometric-gustatory MW (Cytowic, 1982; 1989), such total connection may be observed. The best example of selective correspondences is coloured-lexical synaesthesia, where letters have individual colours. The synaesthete does not have coloured responses to all written characters—only those from the alphabet evoke synaesthetic percepts.

The type of stimulus may also alter the potency of the evoked synaesthetic imagery. A repeatedly documented coloured-hearing and coloured-form synaesthete, DS, reports that the experience of seeing moving photisms in response to music changes with different instruments and musicians (*personal communication*, 1994; Cytowic, 1989; see also Chapter Seven). When listening to a particular singer, the imagery alter from moving in a two-dimensions to three-dimensions, coming towards and away from her point of view. She emphatically stresses that these are her favourite percepts, and are very attention demanding. Another subject, MW, reports that novel stimuli, in his case tastes and smells, evoke the most "vivid and sensuous" sensations (Cytowic, 1989).

2.5.3 Synaesthetic percepts are composed of elemental properties

Reported synaesthetic percepts are almost always composed of elemental properties. Examples of this in visual synaesthetic imagery are clouds, geometric shapes, blobs, and textures. Although these may be combined into florid imagery, they are different in quality to "real-world" objects, such as landscapes and physical objects.

Examples of the elemental nature of imagery experienced by synaesthetes can be found in almost every case study. Most studies of coloured imagery report that the synaesthetes simply assign colour names to the evoking stimulus elements. A coloured-hearing synaesthete examined by Bleuler and Lehmann (1881; cited in Krohn, 1892) reports that "...the noise of respiration is gray. A crackling sound is made up of white points; a tremulous sound is a light bluish gray." The geometric-gustatory synaesthete MW (Cytowic, 1982a) reports feeling shapes and textures in response to stimuli, and describes one such sensory experience as "...a curvature behind which I can reach, and it's very, very smooth. So it must be made of marble or glass..." The imagery is potent but not of a highly specific nature, such as feeling a complex shape of varying texture. The hypothesis that synaesthesia arises from the use of SoCs in development suggests that the scales available to the individual at that early time are rudimentary, and thus the resulting imagery is also rudimentary in nature. Additionally, rudimentary scales may be more useful as SoCs than more complex scales, as mapping onto simpler scales can be performed more lawfully.

The elemental properties of synaesthetic imagery makes it possible to discriminate between synaesthesia and other types of joined imagery experienced in drug-induced states. In the past, these images have been termed as artificial synaesthesia, but differ from synaesthesia in the types of imagery and consistency of imagery experienced by subjects. One study reports a subject, administered with ½ gram of mescaline, who described the imagery of tiny boxing gloves playing a piano keyboard in response to hearing music (Simpson & McKellar, 1955). Other descriptions in this mescaline study outline simple associations. There is a difference between fanciful thinking by subjects and synaesthetic perception; the researchers have assumed that sensory input to a sensory organ resulting in visual imagery defines the experience as synaesthesia. This study does not examine the consistency of synaesthetic responses to the same evoking stimulus—an internal temporal consistency in percept reporting figures in almost all studies as an important criterion for determining the genuineness of a particular synaesthesia.

2.5.4 Synaesthetic associations are idiosyncratic

Synaesthetic percepts are individual for each synaesthete. The earlier mentioned report of Nabokov's coloured alphabet differing from his mother's is an example of this. Several famous coloured-hearing composers show quite different colour assignments with musical tones as well. Messaien, who put specific colour notes in several of his compositions, had colour associations for elements of music more complex than the individual note, such as chords and modal tone centres (Bernard, 1986). However, other composers had individual tone-colour associations as well as associations for higher musical elements. L. B. Castel, a Jesuit of the Eighteenth century, developed one of many colour organs invented over the centuries, and asserted that his colour correspondences were real, and not simply arbitrarily assigned (Marks, 1975). Table 2.1 lists the colour-note correspondences of Scriabin and Rimsky-Korsakov, a 17-year-old coloured-hearing boy, as well as some of the colours for Castel's organ.

As is evident from Table 2.1, there is some agreement for a small number of colours, but no more than would be expected by chance, and nothing that would imply a universal colour-note assignment.

Note	*Scriabin*	*Rimsky-Korsakov*	*17-year-old Boy*	*Castel*
C	Red	White	Bright red	Blue
C-sharp	Violet	Dark & warm	Bright purple	—
D	Yellow	Golden, bright	Green	Green
D-sharp	Steel	Dark	Sky blue (w/green)	—
E	Pearl blue	Dark blue, nocturnal	Blue	Yellow
F	Deep red	Bright green	Yellow	—
F-sharp	Blue	Gray-green	Dark purple	—
G	Orange	Golden brown	Dark red	Red
G-sharp	Purple	Violet	Lavender	—
A	—	Roseate, youthful	White	—
A-sharp	Steel	Dark but strong	Greenish white	—
B	Pearl blue	Steely dark	Normal purple	—

Table 2.1. Colour associations for synaesthetic composers and single musical notes. (Source: Scriabin: Plummer, 1915; Rimsky-Korsakov: Slonimsky, 1955; 17-year-old boy: Rizzo & Esslinger, 1988; Castel: Marks, 1975).

An interesting side note to the colour-note correspondences is the absolute pitch ability that such a synaesthete would have as a result of her associations. If the note *F* always evoked a deep red colour, the perceiver could name the note by employing its synaesthetic percept.

2.5.5 The unidirectional nature of synaesthesia

Synaesthesia is most often unidirectional in nature. Although the primary sense evokes imagery in the secondary sense, few reports exist of the secondary sense evoking a reciprocal sensation in the primary sense. In the case of JR, who does experience bi-directional visual-hearing, the resulting sensory overload makes it difficult for her to be in city environments (*Horizon*, 1995). The other notable case is Luria's famous subject S, whose poly-modal multi-directional synaesthesia may have accounted for his almost limitless memory, as well as his difficulties in leading a "normal" life (Luria, 1968). S's synaesthesia would supply him with constantly changing images accompanying almost all sensory input. The richness of this imagery when listening to human speech made the task of attending to the speaker's meaning, rather than the associations, a demanding task. Both of these conditions are diffuse synaesthesia.

With one-directional synaesthesia, twenty different synaesthesiae are possible if the five sensory modalities are defined as vision, audition, gustation, olfaction, and touch. If other types of experience may be said to evoke synaesthetic responses (*e.g., thinking* of a number evokes a *sense of colour*), or senses are further divided into sub-divisions, then the combinations increase dramatically. Internal pain, such as headache, has been known to evoke visual imagery (DS in Cytowic, 1989; CS; *personal communication*), and fits the criterion of being a primary sensation evoking a secondary one, without fitting neatly into the modality of touch (specifically *haptic* touch). If other such evoking sensations exist, the total number of possible synaesthetic forms is unknown.

2.5.6 Synaesthesia often has a genetic component

There are two important genetic qualities of synaesthesia. First, synaesthesia has been often observed to run in families. The most famous recorded case is that of Vladamir Nabokov, who reports on his synaesthesia in his autobiography *Speak, Memory* (Nabokov, 1966). Both Nabokov's mother and son were synaesthetes: all three experienced coloured alphabets. Other strong evidence for the inheritance of synaesthesia comes from the research of Cytowic (1989), who has documented the genetic pedigree of many synaesthetes. Lundborg (1923; cited in Langfeld, 1926) investigated three generations of a synaesthetic family: the father, three children, and six grandchildren all exhibited some form of synaesthesia. Baron-Cohen et al (1995) have also drawn out family trees for synaesthetic families. Several trees show synaesthesia at three generations, predominately in females. This evidence has led them to assert that synaesthesia may be "...an X linked dominant hereditary condition..." (Baron-Cohen et al, 1995).

The second important genetic quality of synaesthesia is that synaesthetic associations are not inherited. However, researchers have looked for commonalities between family members. Vladimir Nabokov (1966) relates an anecdote of telling his mother that the colours on his alphabet blocks were "...the wrong colour." Nabakov reports on the early manifestation of his synaesthesia and that his mother understood his seemingly odd comment, as she was a coloured-lexical synaesthete herself. Her associations did not match her son's, but she understood the import of his comment nonetheless. Dmitri Nabokov, Vladamir's son, has associations that are idiosyncratic, but attempts to find a basis in his correspondences from a blend of his father and mother's colours (*Horizon*, 1995). This yields possible matches in the alphabet, but no more than would be expected by chance.

The SoC hypothesis of synaesthetic development predicts that the only way mother/son correspondences would be identical is if they used the same SoC for developing an association in an identical fashion. The evidence for the genetic inheritance is still weak, due to the paucity of studies that report other family members with synaesthesia, and lack of common evaluation criteria for synaesthesia. If synaesthesia is genetic, it is probably the tendency to form lasting correspondences between SoCs that is inherited.

2.5.7 Synaesthesia is often associated with higher memory abilities

Synaesthesia has often been observed paired with increased mnemonic abilities. The case of S (Luria, 1968) is the most pronounced of these. S's memory capacity was enormous, with long matrices of digits and serial lists of unrelated words retained for years after a single presentation. In fact, S's pronounced problem was how to forget information. His hypermnesis was attributed, at least in part, to his distinct poly-modal synaesthetic imagery for every element in a presented list. S reported that recall was directly facilitated by the associated imagery, rather than the imagery merely accompanying the information. Interviews often reveal strategies that synaesthetes use for remembering, relying on their synaesthetic associations as a cue for memory retrieval. Cytowic (1989) makes the mnemonic nature of synaesthesia one of the diagnostic criteria (see below, Section 2.6). Synaesthete AL, discussed in Chapters Seven and Eight, demonstrated short-term memory capacity for numbers far outside the normal range. However, requiring the mnemonic use of synaesthetic percepts as a diagnostic criterion should be adopted with caution: other synaesthetes with self-reported high memory abilities fail to demonstrate such abilities when tested with the WMS (see Chapter Eight). Cytowic himself notes that one of his most tested synaesthetes, DS, self-reported having an exceptional memory yet scored in the normal range (101) on the WMS (Cytowic, 1989, pp. 134–135).

2.5.8 Other common reports of synaesthetes

Synaesthetes self-report a variety of abilities and deficiencies that should be mentioned, due to their frequency in recording. They often report being very poor with directional abilities, although studies have failed to show this as an actual deficiency. They report, and researchers have noted, a penchant for unusual or paranormal experiences, including clairvoyance, *déjà vu*, *déjà vecu*, *jamais vu*, *jamais vecu*, precognitive dreams, and telepathy (Cytowic, 1989, pp. 235–236). Cytowic estimates that 17% of the synaesthetes he has worked with report one or more of these types of "unusual experiences" (Cytowic, 1995). The other prevalent quality is the tendency for synaesthetes to be creative. This is especially borne out by the number of famous recorded synaesthetes in many of the artistic disciplines, such as Olivier Messaien, Alexander Scriabin, Nikolai Rimsky-Korsakov, Basho, David Hockney, Charles Baudelaire, Arthur Rimbaud, J.-K. Huysmans, and Vasilly Kandinsky.

2.6 Cytowic's diagnostic criteria for synaesthesia

There is no entry for diagnosing synaesthesia in the DSM-IV, and few methods exist for a formalised diagnosis. This is unsurprising in a community that does not agree on the exact nature of synaesthesia. Attempts have been made in the last decade to create better methods of diagnosis. Still, no one standard exists by which to assess synaesthesia.

Cytowic (1989) proposed a set of five criteria for a DSM-style diagnosis of synaesthesia. Although they are not widely used as yet, they do provide a rigorous framework by which to evaluate potential synaesthetes. These criteria are:

1. *Synaesthesia is involuntary, but elicited.* That is, the perceiver cannot suppress the imagery response to the stimulus. The imagery must also be evoked by a

stimulus; the perceiver cannot imagine the imagery in the same manner at will.

2. *Synaesthesia is projected externally.* The perceiver experiences the imagery as if it is outside of the body, rather than, as Cytowic puts it, "in the mind's eye."
3. *Synaesthetic percepts are durable and discrete.* Synaesthetic imagery should not change during the lifetime of the synaesthete. It is this durability that provides for consistency between tests. The discreteness of the imagery refers to distinct imagery created by different stimuli. Not all noises should evoke images of clouds, for example. The imagery should be as distinct as the stimuli.
4. *Synaesthesia is memorable.* Synaesthesia imagery and percepts are used by the synaesthete as a mnemonic aid. Source stimuli are often recalled by the synaesthetic percept that they evoke. For example, a person might be remembered as "a red name".
5. *Synaesthesia is emotional.* The affective component is strong in synaesthesia: Synaesthetes have non-neutral feelings towards their associations.

These criteria are still quite subjective when diagnosing a possible synaesthete, and require introspection on the part of the synaesthete. The last two of these criteria are contentious. It is possible that all imagery, whether produced voluntarily by a normal subject or involuntarily by a synaesthete can be used mnemonically. It is also likely that all imagery would have some affect tied to it, and that normal subjects do not experience neutral feeling towards any imagery. The first criterion is also problematic, as selective synaesthetes (*i.e.*, coloured-numbers) can experience imagery when "thinking of" a number. This is further documented in Chapters Seven, Eight, and Nine.

2.7 Methods of studying synaesthesia

2.7.1 Questionnaires

2.7.1.1 Wellesley study

Questionnaires have been used in synaesthesia studies since the beginning of the century. A study at Wellesley College collected synaesthetic profiles and statistics over a two-year period from over 200 synaesthetes (Calkins, 1895). A very exhaustive questionnaire was utilised, with questions about most known forms of visual synaesthesia (termed *pseudo-chromaesthesia* in the study). In this study, no defining criteria for synaesthesia was specified and the questionnaire examined a very large array of imagery that might be experienced by subjects in response to categorical stimuli. Despite the exhaustiveness of questioning, the criteria for establishing synaesthesia are loose and optimistic, as the author and researchers were eager to find more cases of synaesthesia to report. The researchers canvassed entire undergraduate classes with the questionnaires, and produced high estimates of incidence. In total, they reported 298 subjects with synaesthesia out of a three-year sample of 979 subjects. One important element of the questionnaire study was the subsequent re-questioning of the reported synaesthetes after a one-year period. Out of this group, they report that, with only *one* exception, all subjects designated as synaesthetes reported the same imagery descriptions when re-questioned. Within-subject reliability on each element was not reported. It is notable here that synaesthesiae being examined all had to do with vision or personality traits as the secondary sense. With the loose standards applied to diagnosis, even more synaesthetes would have most likely been found if other paired sensory experiences were included in the study.

2.7.1.2 Baron-Cohen & Harrison "Test of Genuineness"

In a number of synaesthesia research projects, Baron-Cohen & Harrison have employed a standardised questionnaire (Baron-Cohen, Harrison, Loader, & Rahman, 1993) in tandem with a test they term the *Test of Genuineness*. This questionnaire looked into multiple forms of joined sensation in addition to colour-sense associations. The Test of Genuineness uses a similar technique as the Wellesley study, by re-administering the questionnaire to the synaesthete with no advance warning. The primary criterion in the test is consistency in the reporting of associated imagery or secondary sensations over time. Control subjects were also asked to fill in the same questionnaire, and were tested over time to ascertain the test's effectiveness. In one study, they reported that synaesthetes demonstrated 92.3%

consistency in their responses after periods of a year or more, while control subjects responded with only 37.6% consistency after a single week had elapsed from the first test. Such repeatability lends credence to the constant nature of synaesthetic associations. In fact, repeatability is the one criterion now applied by nearly all researchers as a necessary component of synaesthesia. It is not enough for imagery to occur cross-modally—it must be concrete and temporally constant to quality as true synaesthesia.

2.7.2 Case-studies

Individual in-depth case studies of synaesthesia are somewhat rare, but provide important information about synaesthetic perception. The differing nature of many synaesthesiae make them as widely variant as many brain-damage studies. One study in particular, Luria's assessment of the mnemonist S. (Luria, 1968), seeks to establish how the internal process of synaesthetic experience might account for abilities far exceeding the norm. Cytowic is another researcher who has carried out several case studies. The most thorough of these is a study of an olfactory/gustatory–geometric shape synaesthete, MW (Cytowic & Wood, 1982; Cytowic, 1989). The long association with a single synaesthete allowed Cytowic to carry out pharmacological studies, examining how different drugs altered the nature of MW's synaesthesia.

2.7.3 Brain scan techniques

One of the most interesting advances in synaesthesia research has involved the use of brain scan techniques. Cytowic notes that a common criticism of synaesthesia is its inherent subjectivity. He compares such a purely subjective experience with medically established conditions such as headache and temporal lobe epilepsy (Cytowic, 1989, p. 63). However, recent experiments have demonstrated increased regional cerebral blood flow (rCBF) in the visual associative centres of synaesthetes brains using Positron Emission Tomography (PET) compared to non-synaesthetic controls (Paulesu et al, 1995). This technique provides an objective tool to aid in synaesthetic diagnosis. Harrison & Baron-Cohen (1995) report that preliminary trials using a non-invasive functional magnetic resonance imaging (fMRI) have yielded similar results. The dramatic difference in decreased setup time and more comfortable experimental conditions for fMRI compared to MRI may facilitate the technique's application in cross-modal experiments as well.

2.8 Incidence of Synaesthesia

The researchers at Wellesley (Calkins, 1895) estimated that 50% of all subjects had some fixed form of mental imagery associated with some sense. Additionally, 20% of all subjects are estimated to experience synaesthetic photisms, and 40% to have synaesthetic forms (explained later in Section 2.9.1), although they note that previous estimates by Galton and Flournoy were significantly lower. Vernon (1930) summarised previous literature and cites incidence estimates of coloured hearing from 9% to 43% of the population, and estimates 16% of individuals that filled in his questionnaire had the condition. This is in sharp contrast to the estimate of 1:300,000 by Cytowic (1989). Cytowic first revised this estimate to 1:100,000 (Cytowic, 1993) and then to 1:25,000 and qualifies the estimate by stating that the figure is probably still too low (Cytowic, 1995).

The reason for such variability in incidence estimates may have to do with the way that synaesthetes imagine that others perceive. For most synaesthetes , their associations are simply the way that they have always perceived, and they cannot imagine any other method of perceiving. Synaesthetes often discover that not everyone shares their type of perceptions, and experience disbelief and teasing when discussing their percepts with others (Binet, 1892; Cytowic, 1989).

Synaesthesia has also experienced several periods of increased interest, as mentioned in Section 2.3. Primarily, these times have been at the turn of the century, and within the last 15 years. Binet states that:

> *[Colour hearing] has been repeatedly discussed in the daily papers and literary and scientific reviews; it has been the subject of medical theses and of didactic treatises; it has figured in poetry, in romance, and even in theater; it has given rise to several conventions, the last one of which has just closed in Geneva; physiologists have been pre-occupied with it and have made many experiments concerning it, in their laboratories.* (Binet, 1892)

Despite this high level of interest in the early part of the century, by 1950 very few articles had been published on any form of synaesthesia. In the last decade and a half, research in science and the arts into synaesthesia has been again on the rise. It may be that during those periods when people are more aware of synaesthesia, more recognise that they have the condition, and seek out researchers, and more scientists recognise the condition when they encounter synaesthetes. Cytowic attributes part of the decrease in time of his incidence estimations to communication from self-reported synaesthetes after to the publication of two books on synaesthesia and electronic communication now available, such as Internet, CompuServe, and Prodigy (Cytowic, 1995).

Many researchers estimate that children exhibit synaesthesia much more frequently than adults do. These estimates range from 50% of all children (Révész, 1923) to the more radical proposition that all children are born synaesthetic (Maurer, 1993; Maurer & Mondloch, 1996; Baron-Cohen, 1996). Both suggestions rely on the assumption that synaesthesia is lost in the course of maturation. The latter theory will be discussed later as the *Neonatal Synaesthesia Hypothesis* (see Section 2.10.3).

2.9 Other Synaesthesia-like experiences

There are a number of experiences that are reported in synaesthesia literature that do not involve the crossing of two sensory modality. These experiences fit the description of selective synaesthesia discussed in Section 2.3. Two types of joined sensations, *synaesthetic forms* and *associated personality traits* are outlined here, and discussed further in Chapters Seven and Eight.

2.9.1 Synaesthetic Forms

Synaesthetic forms are imagery that is spatial in quality and are evoked by categorical stimuli. One example of this imagery is number forms, where each number holds a discrete spatial position. When the synaesthete thinks of a number, it is imagined in space along this imaged form. These forms are another example of mapping categorical stimuli onto a SoC in a lawful way. In the process of learning a new concept, such as numbers, the individual may use position in space as an SoC for the evaluation of the numbers.

Synaesthetic forms have been recorded for numbers, seasons, days of the week, months, and other categorical information. Galton recorded many of the earliest cases of synaesthetic forms, along with diagrams of these forms. Figure 2.1 shows two number forms from this collection. The large disparity between the two forms is not merely illustrative: most of the forms are remarkably individual in their layout. One common feature of the different forms is the larger amount of space in the form set aside for the frequently used items in the category (Cytowic, 1989). That is, items that are more important or relevant to an individual occupy more space in the form.

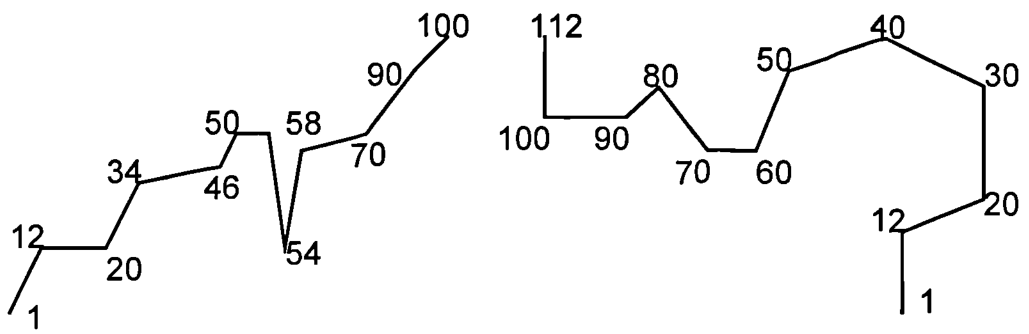


Figure 2.1. Number forms of two subjects. (Galton, 1907).

The similarity that synaesthetic forms have to other synaesthetic imagery is that they are involuntarily evoked, discrete in nature, and durable over the life of the subject (Cytowic, 1989, p. 192). They are also idiosyncratic, with number forms in particular showing a wide variation in their composition. Cytowic (1989) reports that out of the 42 synaesthetic subjects he interviewed, nine (21%) reported number forms as well. He also points out that the number forms of some synaesthetes were not externally projected, and thus do not fulfil his second synaesthesia diagnostic criterion. This thesis will investigate one subject's personal number forms, days of the week form, and seasonal forms in Chapters Seven and Eight.

2.9.2 Personality Characteristics

Personality characteristics are another reported phenomenon that sometimes accompanies synaesthesia. Briefly, these are the sense that some synaesthetes have that a particular element—such as a number, month, musical chord, or day of the week—has a definite gender and personality. These feelings are inexplicable by the perceivers, and they are often aware of how ludicrous such descriptions sound, yet firmly insist that the items "have" those characteristics. One subject insisted that the letters of her initials, **M** and **T**, were masculine, although she was female (Cytowic, 1989; p. 227). Again, these personality assignments have been noted as durable over the person's lifetime, and involuntarily evoked. A study at Stirling demonstrated that subjects agreed upon which adjective from a pair representing a continuum of a human personality traits (*i.e., Good/Evil, Happy/Sad*) was most appropriate for short passages of music as stimuli (Watt, Gerstley, & Ash, *in preparation*). This was further extended by a study that demonstrated that the same types of adjectives were matched by subjects to monophonic musical sequences of three or more beats (Watt & Ash, 1997). Further studies along these lines will be presented in Chapter Five.

2.10 Theories of synaesthesia

As the question of precisely what constitutes synaesthesia has been in dispute, it is inevitable that the theories explaining the cause of synaesthesia have been quite varied. The major theories will be reviewed here by groups.

2.10.1 Undifferentiated neuronal activity

The Undifferentiated Neuronal Activity theory first became popular in the nineteenth century. This theory postulates that synaesthesia is the result of "sensory incontinence." Sensory information in synaesthetes is believed to be processed diffusely, rather than in any single modality. This theory likened synaesthesia to synkinetic movement in infants; a result of an immature nervous system (Cytowic, 1989, p. 67). Although Cytowic dismisses these theories out of hand, the notion that sensory incontinence may be present at some point of development is important in the Neonatal Synaesthesia hypothesis (see Section 2.10.3), which begins with the assumption that newborns experience sensory information initially in a diffuse manner, with sensory modularisation normally developing and replacing synaesthetic perception. The brain scan evidence from Paulesu et al (1995; see Section 2.7.3) does demonstrate activity in several sensory centres in response to single modality input, which would be predicted with this theory.

The difficulty with such a theory is that it cannot predict why synaesthetic responses develop for categorical stimuli. It is improbable that only sounds that are musical in nature should generate one-to-one mappings with colours when the tonal scale must first be learned through experience. It also cannot explain categorical stimuli producing imagery when the specific origin of the association can be identified (*e.g.*, coloured fingering charts representing the same notes that a synaesthete experiences in response to that note).

2.10.2 The Cross-Modal Transfer hypothesis

The Cross-Modal Transfer hypothesis (CMT) and Neonatal Synaesthesia hypothesis (NS; see Section 2.10.3) both have a basis in neonatal experimentation. Each states that synaesthesia develops when a poly-sensory association which occurs as a normal part of cognitive development becomes fixed and permanent. The CMT hypothesis posits that neonates "...are capable of using and storing surprisingly abstract information about objects in their world." (Meltzoff & Bornton, 1979; p. 404). When an object is explored through one modality, abstract qualities of that object are retained in memory. When the object is explored through another modality, the object is recognised as being the same if the resulting abstract representation of the object is sufficiently similar to the first representation (Baron-Cohen, 1996).

A study by Meltzoff & Bornton (1979) examined how babies recognise a visually presented object that had been previously explored orally. Briefly overviewed in the last chapter, the results showed that babies looked at a picture of the pacifier, either smooth or nubby, that they had previously explored. Maurer (1993) describes the abstract representations of the pacifiers as having continuous and discontinuous energy, respectively. For example, the smooth pacifier stimulates continuous energy in the sensory system regardless of input modality, and would then subsequently be recognised when encountered independent of modality. The results in light of the infants' very young age (29 days) provide compelling evidence for an innate ability to recognise objects across modalities. The studies of Lewkowicz & Turkewitz (1980) also show similar cross-modal matching of audio and visual properties of stimuli by infants.

The CMT hypothesis suggests that synaesthesia is a result of a subset of abstract representations across sensory modalities becoming permanent. This theory does not, however, predict how abstract representations produce imagery in response to sensory input. In addition, it does not predict the unidirectional nature of synaesthesia: if the abstract representation of one sense produces imagery, nothing in the model states why the imagery should not have a reciprocal when stimulation is received in the other sense.

2.10.3 The Neonatal Synaesthesia hypothesis

Outlined in the last chapter, the Neonatal Synaesthesia hypothesis is built upon the ideas of undifferentiated neuronal activity and the Cross-Modal Transfer hypothesis (Section 2.10.2). The difference between the CMT and NS hypotheses is subtle. The CMT hypothesis holds that sensory input from an object is represented abstractly and the comparison of two different representations indicates whether an encountered object is novel or familiar. The NS hypothesis states that the neuronal activation pattern produced by sensory input may be the same regardless of input modality, and a familiar activation pattern is recognised while novel activation patterns are not.

The NS hypothesis asserts that an infant's senses are undifferentiated at birth, and thus sensory input is processed diffusely in the cortex. The hypothesis holds that we all start out life as synaesthetes. Maurer (1993) hypothesises that infants behaviour is governed by two principles: "(1) keep the sum of energy entering all sensory channels within an optimal range and (2) when the sum is at an appropriate level, attend to the familiar patterning of energy *regardless of the modality of origin* until a schema is well-formed, then search out a novel pattern of energy." (p. 112; author's italics).

Maurer (1993) explains the Meltzoff & Bornton (1979) results by suggesting that visually smooth and tactilely smooth pacifiers evoke continuous patterns of energy in the sensory system, while nubby pacifiers produce discontinuous patterns of energy (Maurer, 1993). In an immature system, this follows along the lines of Mark's *Doctrine of Common Psychophysical Properties* (Marks, 1978). It is important to note that infants might abstractly represent continuous and discontinuous energy in a similar manner described in the CMT hypothesis, and thus not differ greatly from that model.

If diffuse activity accounts for the cross-modal associations, then cross-modal transfer should decrease as the sensory system becomes modularised. Maurer (1993) attributes failures to replicate the Meltzoff & Bornton (1979) results with two- to four-week olds to levels of stimulation too high for the infant. Streri (1987) and Streri & Pecheux (1986) observed the predicted decrease in cross-modal transfer in four- to five-month old infants in contrast to two-month old transfer performance.

Although the NS hypothesis describes the process of neonatal cross-modal transfer, it seems incorrect to regard recognition of stimuli across modalities as the same type of joined sensation as mature synaesthesia. Synaesthesia in almost all studies refers to a joined sensation: the NS hypotheses describes a process of amodal perception rather than imagery evoked by a stimulus. The hypothesis also does not suggest how the neonatal synaesthesia develops into adult synaesthesia. Presumably it results from anomalous development resulting in incomplete modality differentiation. There is no evidence that the types of the neonatal synaesthesia experienced by neonates develops into the types of synaesthesia recorded in mature individuals.

The NS hypothesis also cannot account for adult cross-modal transfer, as Maurer points out. Maurer asserts that "As the cortex matures, the early apparent cross-modal transfer should decrease, followed by the development of a more analytic cross-modal transfer." (Maurer, 1993). Thus, in the course of normal development, schemata are formed initially using the NS model until the child is approximately 4 months old. After this period, then schemata subsequently form according to another unspecified process.

The NS hypothesis also cannot account for synaesthesiae which are limited in their associations: selective synaesthesiae. The paradigm is better suited to diffuse synaesthesiae with its wide modality crossings, such as having all sounds evoke photisms. The problem becomes further exacerbated when considering the language based synaesthesiae. Written language is not encountered until the infant is well past the hypothesised end of neonatal synaesthesia.

These criticisms should shed no doubt upon the actual existence of neonatal synaesthesia—only on its basis for all adult synaesthesiae. As mentioned earlier, several different types of synaesthesia may exist, common in their key attribute of sensory crossing, but possibly differing in their geneses and underlying processes. The more dramatic forms of synaesthesia—such as the poly-modal multi-directional imagery of S (Luria, 1968), the bi-directional audio-visual imagery of JR (Baron-Cohen, 1996; *Horizon*, 1994), the geometric imagery in response to all tastes for MW (Cytowic and Wood, 1982b; Cytowic, 1989): each may result from the incomplete modularisation of the sensory system according to the NS hypothesis. Other synaesthetic conditions with imagery that is produced from categorical stimuli—coloured-music, coloured-language and numbers, number forms, and coloured seasons—may result from permanence in connections formed according to either the CMT paradigm or a product of what Baron-Cohen et al (1993) termed a "breakdown in modularity." This *breakdown* was speculated to be the result of modular connections between sensory centres that have formed, grown, or not died out during development. The resulting inter-connectivity leads to a modularity breakdown, which may produce synaesthetic imagery.

One remaining problem with the NS hypothesis is its inability to explain the unidirectional nature of most synaesthesiae. The possible existence of more than one origin for synaesthesia may again explain this, but the rarity of bi-directional synaesthesia, even amongst the more dramatic forms of synaesthesia, makes this unlikely. The evolution from diffuse to modularised senses may incompletely form so that information still flows in one direction but not the other. Further research with the pacifier task to determine if cross-modal transfer would take place from the visual task to the tactile would answer part of this objection.

2.10.4 Limbic theory of synaesthesia

Cytowic (1982b, 1989) proposed a theory that localises the neural structure responsible for synaesthesia as the limbic system. He describes his theory as a *linkage theory*: in his terms he regards synaesthesia as a process of *polymodal combination* rather than the other theories that use *polymodal abstraction*. According to this distinction, polymodal combination is an additive process where the eliciting stimulus and its accompanying imagery are an experience where neither "loses their individual identities." (Cytowic, 1989, p. 72); polymodal abstraction is proposed as a subtractive process that extracts common characteristics of each experience as the mediator for synaesthetic imagery. He rejects polymodal abstraction on the grounds that the process produces an "abstract *perceptual* residue" rather than the more complete sensation that synaesthetes report. This objection speaks only to the creation of the association, however, rather than to the nature of the synaesthetic imagery.

Cytowic suggests that the limbic system, and specifically the hippocampus, plays the role of linking together sensations, creating the polymodal combination:

It is possible to bring information that was processed in geographically separate parts of the brain together as signals in a unique structure that knows about the internal milieu and about the fundamental goals of the organism as a biological entity. That structure can also respond back to the cortices that initiated the circuit and further act on autonomic structures that govern the internal milieu. One has, then, a fundamental yet ideal device that brings information together in the context of how the organism is what it wants to be. It is on this organ that I believe the expression of synaesthesia depends. (Cytowic, 1989, p. 175).

Cytowic rejects a cortico-cortical model of synaesthesia (see Section 2.10.5) on the grounds that such a model would involve the abstraction of sensory qualities, such as Aristotelian common sensibles. He postulates that if this was the case, separate evaluation on the semantic differential task (Osgood et al, 1957) of an evoking stimulus and its associated image should reveal identical differential ratings. The failure of geometric-gustatory synaesthete MW to provide similar semantic differential ratings for gustatory sensations and their associated imagery supports this assertion. Cytowic views synaesthesia as an additive process, where the sum sensation has qualities from both involved senses *as well as* additional elements—primarily affective. For this reason, he searched out a neural structure that would combine sensory information. The mnemonic capabilities often observed in tandem with synaesthesia also led to the limbic hypothesis.

2.10.4.1 Neural evidence for the Limbic hypothesis

Cytowic has based most of the evidence for the Limbic hypothesis on the single case study of MW. He was able to explore MW's synaesthesia with pharmacological studies and brain scans of regional cerebral blood flow (rCBF) using a radioactive xenon inhalation method (Cytowic, 1996; 1993, 1989). With this technique, a decrease in synaesthetic imagery potency was observed when the cortex was stimulated with dextroamphetamine, as well as a large increase in synaesthetic potency when the cortex was depressed or rendered ischemic through the use of ethanol and amyl nitrate, correspondingly. He also observed dramatically reduced cerebral blood flow while MW was experiencing synaesthesia—so marked that Cytowic states that a normal subject with the same blood flow would have been "...blind, paralyzed, or showing other conventional signs of a lesion," yet MW's thinking was reportedly unimpaired in neurological assessment, and no lesion could be located (Cytowic, 1996; 1989).

It is notable that Cytowic did not actually observe increased limbic activity. Instead, he infers its importance in synaesthetic perception due to its proposed importance in memory and emotions. Maurer & Mondlach (1996) also propose that the limbic system may be a centre for synaesthesia—in their model, it may account for neonatal synaesthesia. They note that the cortex of the new-born is "hardly functioning" (Johnson, in press; cited in Maurer & Mondlach, 1996), while the limbic system is believed to be functional at birth.

Limbic combination may, like the NS hypothesis, account for the diffuse synaesthesiae, but is inadequate to explain selective synaesthesia. It suffers from the same objection as the NS hypothesis in that it does not predict the why synaesthetic imagery can be produced by categorical stimuli.

2.10.5 Cortico-cortical Connection hypothesis

The final hypothesis to be reviewed is the Cortico-cortical connection hypothesis, introduced by Baron-Cohen, Harrison, Goldstein, & Wyke (1993). This model attributes synaesthesia to connections present between different cortical sensory areas, causing activity in the associated modality when the primary modality is stimulated. This is essentially a neurological equivalent of the Cross-modal Transfer hypothesis (Section 2.10.2). The most compelling support for this model comes from the Positron Emission Tomography (PET) experiments comparing rCBF in both synaesthetes and controls (Paulesu et al, 1995).

This study stands out for its sample size with balanced controls, consistent criteria for evaluating synaesthesia, and objective evaluative technique. Six chromatic-graphemic synaesthetes and six controls matched for gender and handedness were presented with single tones and single spoken words and asked to respond with a finger tap after hearing a stimulus. PET results revealed significant differences in rCBF for synaesthetes when compared to controls. Significant increases in rCBF for synaesthetes were observed in the left posterior inferior cortex, middle frontal gyrus, and right insula; decreases were observed in the left insula and left lingual gyrus. A review of neuropsychological studies in the paper ties the increases in rCBF to various cortical areas hypothesised to be involved in colour perception. The decrease in blood flow proved more problematic to explain, with only speculation into a potential "shutting off" of areas not necessary for the synaesthete's task compared to controls. Interestingly, they did not observe significant differences in the lower-level visual areas of synaesthetes (V1, V2, and V4).

These results are quite convincing, and may provide a method to objectively observe and diagnose synaesthesia. Even so, a few problems with using this method to evaluate all synaesthesiae exist. This experiment only examined one of the many forms of synaesthesia, albeit a relatively common one. The synaesthetes in the study were chromatic-graphemic, with colours associated to a word's initial letter, rather than the initial sound (chromatic-phonemic) or an overall colour-word association (chromatic-lexical) (see Section 2.10.2; and Baron-Cohen et al., 1993). Similar tests with these other types of coloured language, as well as other synaesthesiae with visual imagery, would provide an even fuller picture of the neurophysiological centres common to synaesthesia. Examinations of less selective synaesthesia in which an entire sensory modality has associated synaesthetic percepts may also provide quite different patterns of rCBF. Indeed, just such a difference between types of synaesthesiae may account for the large discrepancy of results from Paulesu et al and Cytowic's xenon inhalation experiment with MW (Cytowic, 1989). It is especially important to note that the Paulesu experiment did not observe any increase or significant change in limbic system activity for synaesthetes, casting doubt on the Limbic hypothesis' ability to account for chromatic-graphemic synaesthesia. Indeed, Cytowic's brain scan study was limited in the types of brain activity that could be observed, and could not demonstrate *increased* limbic activity.

One objection to raise here is that activity in visual areas may be accounted for by mental processes other than synaesthesia. Murata, Cramer, and Bach-y-Rita (1965) found that areas in a cat's visual cortex responded to sound and skin pricks as well as light. More pertinent to human perception, Zatorre, Evans, and Meyer (1994) observed an increase in rCBF in the occipital lobe when subjects with little formal musical training listened to structured music. The activation in visual areas remains unexplained, and Zatorre reports further observing increased rCBF in putatively visual brain areas with speech as well as olfactory perception (1997, *personal communication*).

In the Paulesu et al (1995) study, no visual activity in controls was observed. However, control subjects were not instructed to perform any visualisation during the experiment. The synaesthetic subjects knew that visualisation was expected of them due to their knowledge of synaesthesia and the documenting of their specific associations with the researchers (*e.g.*, repeatedly taking the "Test of Genuineness"). Repeated observations of different rCBF patterns for synaesthetes and controls when controls performed a visualisation task while being presented with auditory stimuli would strengthen the Paulesu et al findings.

Further studies using brain scan techniques are possible, and would aid in localizing potential centres of synaesthetic interaction. The recent successes of Harrison & Baron-Cohen to replicate the PET results using non-invasive functional Magnetic Resonance Imaging (fMRI) with its reduced set-up time of minutes rather than hours should facilitate such further studies, and perhaps allow for the inclusion of children in the brain scan research.

As the PET scans and fMRI scans both reveal an increase in the number of cortical areas active during the most common of synaesthetic perceptions, the Cortico-cortical model best accounts for the selective synaesthesia. It is hypothesised that the continued cross-modal use of SoCs by synaesthetes results in the creation of cortico-cortical connections by the process of *modularity breakdown* suggested by Baron-

Cohen et al (1993), while connections between sensory areas form, grow, or do not die out in the normal course of development. The Cortico-cortical physiological model is then the result of the psychological model of using cross-modal SoCs.

This combined Cortico-cortical/SoC model accounts for the existence of virtually any sensory crossing. The prominence of vision in perception predicts the majority of synaesthesiae which have visual imagery as their secondary sense, as more visual SoCs are available for evaluating novel stimuli, and may have the salience to more comparative tasks than any other sensory modality.

2.11 Structure of this thesis

The experimental section of this thesis will begin by extending the cross-modal findings of earlier experiments by using monochromatic visual stimuli and asking individuals to match them to more complex auditory stimuli, specifically basic musical constructs in Chapter Three. The experiments will be analyzed in the light of the SoC hypothesis, examining which SoCs subjects use to form their matches between stimulus sets. Chapter Four will further extend these results by making the visual stimuli more complex and ecologically valid for comparing cross-modal results to synaesthesia: the auditory stimuli will be identical, but the visual stimuli with now be of a wider spectral range, while still attempting to control for other visual factors. Chapter Five extends the results even further, by using even more complex musical stimuli. Chapter Six pushes the cross-modal task to an extreme, by investigating how subjects perform when asked to rate a complex musical stimulus (an entire song) to a set of adjective scales. Finally, Chapters Seven, Eight, and Nine examine the associations of a small number of actual synaesthetes, as well as some synaesthetes' performances in the same cross-modal tasks from the earlier chapters, and how the SoC hypothesis can explain their synaesthesiae and performance in experiments.

3 Cross-Modal Matching of Simple Musical Constructs and Monochromatic Visual Stimuli

3.1 Overview

Cross-modal matching between musical stimuli and circles of varying grey levels are examined. An initial condition of matching grey-levels with individual musical tones demonstrated regular matching of brighter circles with higher tones. The results also reveal a small percentage of subjects performing with opposite direction of matching. This technique was then extended to examine matching with two-note temporal intervals and three- and four-note chords. Matching brightness with different aspects of intervals and chords were observed between subjects, demonstrating that cross-modal matching can be used to indirectly observe the relative saliency of various musical dimensions for that subject. The effect of a preceding tone on brightness matching with the stimulus tone weakens a common cross-modal auditory-visual space hypothesis. Brightness ratings of F-chord variants provide a subjective ordering of chords, and demonstrate that some quality of inverted chords are aligned with brighter images for most subjects, supporting the hypothesis that they are perceived as higher in pitch-class than root-position chords.

3.2 Introduction

This chapter will examine individual perception of simple musical stimuli, such as individual tones, two-note intervals, and chords, and the manner in which these dimensions can be aligned with an artificially presented dimension of varying brightness. As outlined in the Chapters One and Two, a general correspondence between brightness and pitch height has been hypothesised for centuries. Experiments have revealed regular cross-modal matching between increasing pitch height and increasing visual brightness (Marks, 1974). Similar experiments introducing changes in volume demonstrate that brightness can be aligned to volume (changes) in addition to pitch height. Subjects thus use the single *Scale of Comparison*

(SoC) of varying brightness to align a combination of attributes inherent in the auditory stimulus. The two percepts need not share a common evaluative area, a supra-sensory area, in order to explain such matching. Rather, a dynamic process of forming alignments between SoCs is hypothesised to explain cross-modal correspondences. Subjects can subsequently use such alignments as a temporary schema for perception and categorisation.

Using this paradigm, the consistency of such alignments between subjects is of great interest. If SoCs from several sensory modalities are aligned in the same direction by a significant majority of individuals, then some underlying connecting quality of the sensory experiences must be held in common across the population. Some possible connecting qualities may include magnitude, affective responses, or semantic association for each percept.

Cross-modal matching tasks which use varying visual brightness as one half of the stimulus set provide a method of examining aspects of other sensory continua. Stevens & Guirao (1963) used a similar paradigm in their examination of cross-modal matching of line length with loudness and brightness. In their words, line length judgments were useful not as judgments in and of themselves, but:

> *...it is sometimes possible with an easier task to discern features that are hard to detect in judgments of magnitude on more difficult continua. In addition there is the attractive possibility that apparent length may be used as the standard or criterion against which to scale other continua. (Stevens & Guirao, 1963)*

Similar logic may then lead to the use of apparent brightness as another standard by which to examine other continua. Again, this does not mean to say that the stimuli being compared *are* brighter or longer, but that instead they align regularly to this SoC.

Examining cross-modal matchings may also provide some insight into synaesthetic associations experienced by visual-auditory (coloured-hearing) synaesthetes (and indeed all synaesthesiae). Although synaesthetic associations are idiosyncratic in nature, synaesthetes often report imagery where brightness plays an important distinguishing role: when attempting to describe their imagery, they often turn first to descriptions of the photism's *brightness*. Regular cross-modal correspondences in normal subjects may then provide evidence pointing to the key factors involved in forming synaesthetic correspondences.

3.3 Experiment One: Cross-modal Matching of Monochromatic Colours and Single Musical Tones

3.3.1 Aims and Introduction

Experiment One attempted to replicate earlier findings of cross-modal associations between pitch height and colour brightness by using different experimental techniques. Each of these techniques needed to be evaluated before they could be used in subsequent experiments. The first of these techniques was the use of a two-alternative forced choice (2AFC) paradigm to examine cross-modal relationships. The second consideration was the validity and effectiveness of using a VDU for presentation of colour choices. This presentation method was proposed to allow for greater experimenter control over the visual stimuli. The use of the NeXT computer's VDU allows for the precise production of over 1,000,000 colours, versus the limitations of hand-produced stimuli. Previous cross-modal experiments used a very limited computer produced colour set (Melara, 1989a; Melara, 1989b), Munsell Colour Chips, or incandescent white light (Marks, 1974). Thirdly, the experiment introduced the use of a more complex tonal stimulus: complex waveforms were used, rather than the relatively simple sine waves used in many cross-modal studies (Holmgren, Arnoult, & Manning, 1966; Marks, 1975; Schneider & Bissett, 1981; Melara, 1989a; Melara, 1989b). These complex waveforms are more representative of the types of tones typically found in Western music, lending more ecological validity to the stimuli.

3.3.2 Method

3.3.2.1 Subjects

The subjects consisted of 39 undergraduate students (10 male, 29 female) drawn from the psychology courses at the University of Stirling. Four of the subjects were musically trained, which was defined as having more than one year of musical training or practice on a musical instrument or in voice. All subjects were given credit for participation, partially fulfilling a degree requirement.

3.3.2.2 Equipment

Visual stimuli were presented on a NeXTstation TurboColor computer, equipped with a 19" VDU. Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a Yamaha DX-7 Synthesizer using its standard *Harp-Flute* patch. The synthesizer's output was synchronised with the NeXT computer by means of a standard Midi connection. The subject was seated at a comfortable viewing distance from the VDU (50 centimetres), which subtended a 29° visual angle. The subject provided input to the NeXT by means of the two-button mouse.

3.3.2.3 Stimuli

3.3.2.3.1 Audio Stimuli

The twelve Western tonal notes, covering a range of over three octaves, ordered as: F_3, A_3, Cs_4, E_4, G_4, Bf_4, C_5, Ef_5, Af_5, B_5, D_6, Fs_6. Figure 3.1 shows the tone stimuli in musical notation. The stimuli were presented binaurally through headphones at an approximate volume of 62 dB.

Figure 3.1. Single tone stimuli for Experiment One.

3.3.2.3.2 Visual Stimuli

Subjects were shown two grey circles simultaneously with the auditory tone. The circles varied in grey level from 0.1 (near black) to 0.9 (near white), each two separated by an interval of 0.4. Table 3.1 shows the brightness of each of these grey levels measured in candelas per square meter (cd/m^2), with accompanying coordinates for the grey colour in CIE 1924 Standard Observer colour space. The circles were displayed side by side on the screen as shown in Figure 3.2. Each circle subtended a 13° visual angle.

Grey Level	*Candelas per square metre*	*CIE X*	*CIE Y*
0.1	18.2	0.361	0.375
0.2	26.6	0.337	0.355
0.3	36.0	0.322	0.343
0.4	45.3	0.315	0.337
0.5	52.4	0.308	0.333
0.6	62.9	0.301	0.327
0.7	71.5	0.296	0.323
0.8	82.8	0.295	0.322
0.9	88.3	0.295	0.323

Table 3.1. Candelas per square metre and (X, Y) coordinates in CIE colour space for each of the nine grey levels in Experiment One.

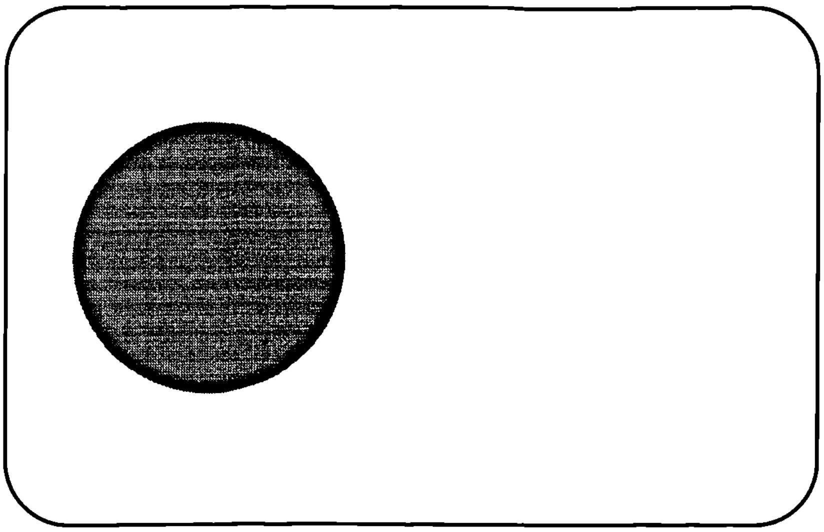


Figure 3.2. Sample monochromatic stimuli for Experiment One.

3.3.2.4 Procedure

The subjects were told that they would hear a tone through the headphones while two circles would simultaneously appear on the computer monitor. They were instructed "...to select the circle that you feel is most like the tone. If you feel that the one on the left is more like the tone, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of tone and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The subjects were played each of the 12 tones four times, making a total of 48 trials. These tones were presented in a random order. The program randomly selected a grey level for one of the two circles that ranged in grey value for 0.1 to 0.5. The selection of the one circle determined the grey level of the second circle, as the second circle was offset by a constant difference of +0.4 for discriminability, as described above. The left or right positioning of the two circles was randomised on each trial. The subject was thus provided with forty-eight separate trials of tone and circle pairs. The tones and circles were presented simultaneously. The duration of the tone was 3 seconds. The circles remained visible until the subject gave their response. If the requested, any single trial could be repeated so that they could hear the tone again. At the conclusion of the last trial, the screen went blank, and a "Thank You" message appeared. At this point, the subject proceeded either on to Experiment Two or Experiment Three, selected in a random fashion.

3.3.3 Results

Figure 3.3 shows the results of Experiment One. Each data point along the abscissa represents one of the auditory stimulus levels. In this experiment, those levels were the pitch heights of the auditory tone. The small box at each point shows the mean score for each stimulus. Scores for each stimulus are computed by first determining whether the subject chose the brighter (closer to white) or darker (closer to black) circle for an individual trial, a bright selection scored as 1, and a dark selection as 0. The sum score for each pitch level is then computed for each subject at each pitch level, yielding 12 separate scores per subject. Each pitch score thus ranges from 0, indicating that the darker circle was chosen all times at the pitch level, to 4, indicating the brighter circle was chosen each time. The whiskers at each data point plot the standard deviation of scores for that data point, while the box around the mean represents the standard error for that mean. In the graphs, the standard deviation size may be relatively wide, indicating that subjects provided a wider range of scores for each stimulus; however, the standard errors are very small, due to the large number of samples for each data point.

A one-way repeated measures ANOVA examining scores for each pitch height reveals highly significant differences between the means ($F(11, 418) = 30.36$, $p < .001$). A correlation of scores and pitch height yields a highly significant positive trend (Pearson's $\rho = 0.63, p < .001$).

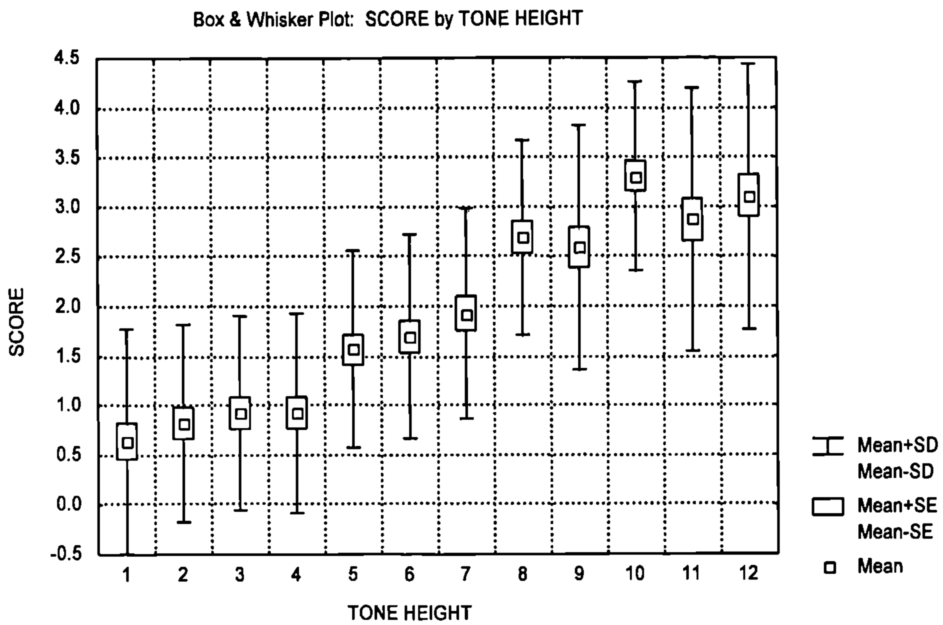


Figure 3.3. Mean scores, standard deviations, and standard errors for all subjects ordered by stimulus tone height.

Individual correlations were run on each subject's ratings to ascertain direction of association. Table 3.2 shows the correlation coefficients for each subject's ratings in columns, as well as the significance of these correlations. Figure 3.4 shows a histogram of the correlation slopes found for all subjects. Regression slopes for a majority of subjects (56%) exceed +0.3. Twenty-seven subjects show a highly significant positive correlation between pitch height and brightness of circle chosen (Pearson's $\rho > 0.701$; $p < .01$; 69% of all subjects), three subjects show a significant positive correlation (Pearson's $\rho > .571$; $p < .05$; 8% of all subjects), and 2 subjects show a highly significant negative correlation (Pearson's $\rho < -0.701$; $p < .01$; 5% of a subjects). Only seven subjects (18%) do not demonstrate an observable alignment between pitch height and circle grey level. No significant differences between male and female performances are observed.

Subject	ρ	*p-value*	*Subject*	ρ	*p-value*
1	0.777	0.004	*21*	0.907	0.000
2	0.916	0.000	*22*	0.907	0.001
3	−0.039	0.906	*23*	0.924	0.000
4	0.955	0.001	*24*	0.716	0.010
5	0.740	0.006	*25*	0.860	0.001
6	0.905	0.000	*26*	0.925	0.001
7	0.695	0.013	*27*	0.809	0.002
8	0.741	0.006	*28*	−0.171	0.594
9	0.825	0.002	*29*	0.652	0.022
10	0.867	0.001	*30*	0.813	0.001
11	0.725	0.008	*31*	0.927	0.000
12	0.925	0.000	*32*	−0.736	0.007
13	−0.143	0.657	*33*	0.769	0.004
14	0.894	0.000	*34*	0.493	0.104
15	0.832	0.001	*35*	0.879	0.001
16	0.336	0.286	*36*	0.874	0.000
17	0.429	0.165	*37*	−0.911	0.000
18	0.921	0.000	*38*	0.642	0.001
19	0.655	0.021	*39*	0.722	0.008
20	0.183	0.001			

Table 3.2. Correlation coefficients for individual subject performance in Experiment One

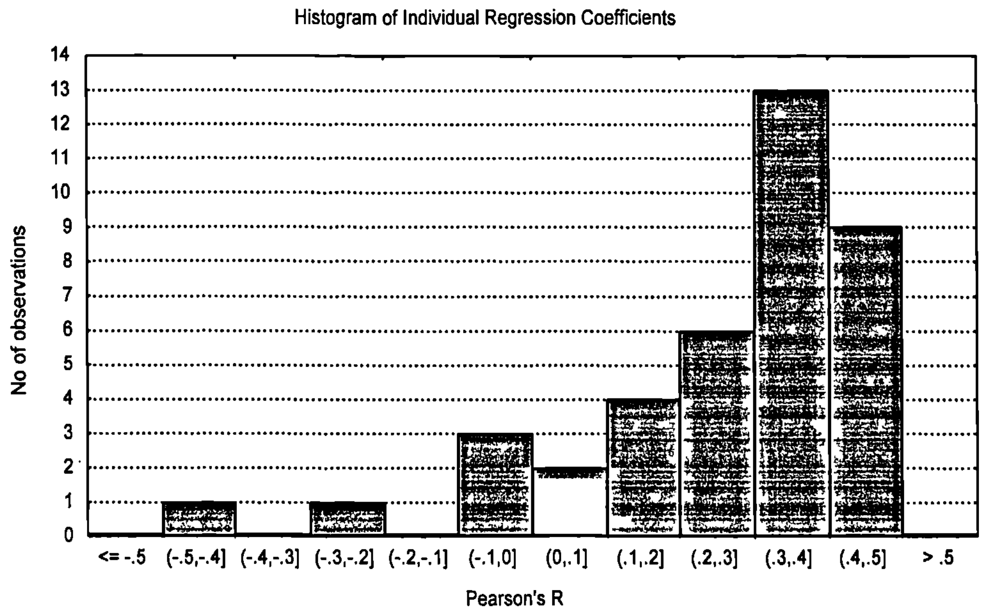


Figure 3.4. Histogram of individual correlation coefficients for all subjects. Coefficients computed for score and pitch height. Positive correlation indicates matching of brighter circles with higher tone; negative correlations are brighter circles with lower notes. Correlations greater than 0.2 are significant (p < 0.05) and correlations greater than 0.3 are highly significant (p < 0.001).

3.3.4 Discussion

All of the aims of Experiment One were satisfactorily achieved. The use of the 2AFC paradigm did succeed in revealing regular cross-modal matching performance for subjects. The use of musical tones more complex than simple sine waves is also supported from the regular matching performance. Also, presentation and response recording using the NeXT and its VDU was shown to be a valid technique.

The results replicate previous studies by demonstrating regular cross-modal matching of stimuli with increasing auditory pitch height and increasing visual brightness. Generalised across the subjects, Figure 3.5 demonstrates how strong and regular the tendency to match pitch height and brightness together.

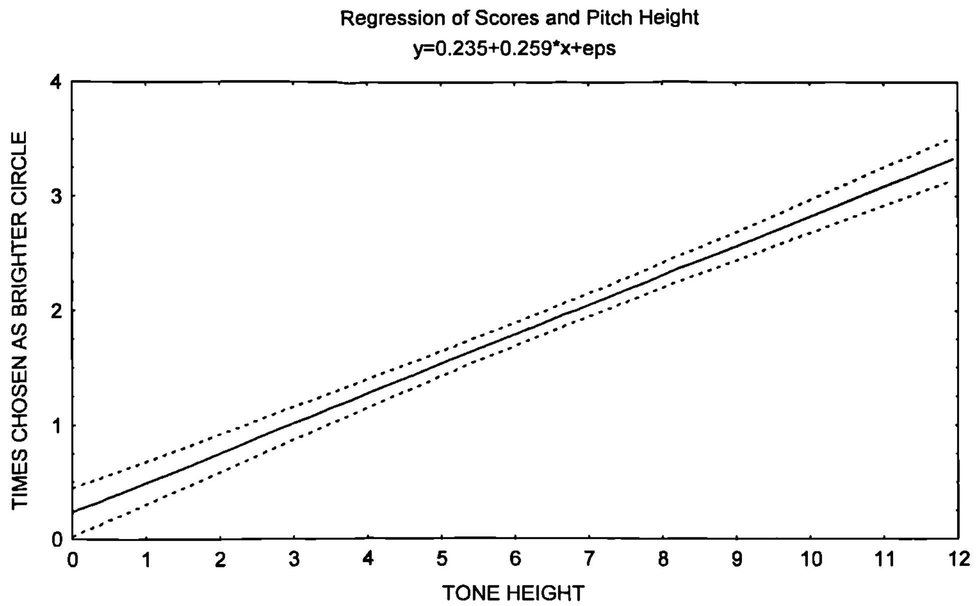


Figure 3.5. Regression of subject scores at each level and pitch height of the tone.

The individual subject regression analysis reveals cross-modal associations that may have been present but unreported in previous studies: some subjects respond in an opposite direction from the majority, but still respond in a consistent manner. Subjects 32 and 37 each matched darker circles with higher tones and brighter circles with lower tones. This matching performance would not have been apparent without the individual analyses, as the strength of the remaining subjects' associations would obscure it. It is possible that such cross-modal matching performances were present and never reported in earlier studies.

The matching performance of Subjects 32 and 37 may seem counter-intuitive; however, one comment that a chromaesthete made during the debriefing of a later experiment may help explain this. When told that the experiment was attempting to determine if there was a relationship between the pitch height and brightness, and told that this would mean that the blacker circles were thus lower tones, she commented, "That's ridiculous. That black circle is much 'brighter' than this grey one." While she still reported a similar correspondence, her sense of what constituted a "bright" attribute of a circle differed from the measured attribute. It is possible that both the chromaesthete and Subjects 32 and 37 were matching auditory tones with colour contrast, as the black circles had the highest contrast to the white background they were presented upon.

It is important to note that effects due to memory for the stimulus tones may be present in the comparisons. The tones were presented for three seconds and then ceased while the subject could continue looking at the circles to make their choice. The possible memory effect should be small, as most subjects responded within the first three seconds of the trial, and had the option to hear the tone repeated upon request.

Strong individual differences in both directions of cross-modal matching do not support the hypothesis of a suprasensory representation of brightness (Hornbostel, 1931). Instead, cross-modal matching may be accounted for by visual brightness or contrast functioning as effective SoCs for alignment with pitch height. Matching performance is dictated by the SoCs chosen, and the direction these SoCs are matched. Aligning the SoC of contrast so that increasing contrast was aligned with increasing pitch height *or* the SoC of visual brightness so that decreasing contrast aligned with increasing pitch height both produce matchings opposite from the sample trend.

3.4 Experiment Two: Cross-modal Matching of Monochromatic Colours and Two-Note Phrases

3.4.1 Aims and Introduction

Experiment Two employed the techniques proposed in Experiment One and explored cross-modal matching performance with stimuli that were more musical. As was mentioned earlier, previous cross-modal auditory pitch and visual brightness matching experiments used single tones as auditory stimuli, and these tones were most often simple sine waves, sometimes called pure tones (Marks, 1974; Melara, 1989a; Melara, 1989b). Experiment One demonstrated that complex tones, or musical notes, used as stimuli in a simple monochrome matching task still reveal a pattern of subjects aligning rising pitch height with increasing visual brightness. Experiment Two was performed to examine how cross-modal matching performance would change when the auditory stimuli were increased in complexity. To this end, the experiment examined cross-modal matching of visual brightness with a simple musical construct: a two-note temporal phrase.

The experiment used only three target notes as stimuli to be matched with the grey circles. On any one trial, the target note was preceded by a second note, which made it possible to begin examining musical context effects on cross-modal matching. The target notes were preceded by notes higher and lower, which may reveal an effect from the direction of approach on the matching of visual brightness and musical note. The size of the intervallic leap was also varied, to potentially reveal effects from the size of the intervallic leap. Finally, the types of intervals formed by the two notes were varied, so that each of the twelve standard intervals formed from one octave of a western chromatic scale are presented. This may reveal effects from learned musical qualities, such as interval consonance.

3.4.2 Method

3.4.2.1 Subjects

The subjects consisted of the same 39 individuals who participated in Experiment One.

3.4.2.2 Equipment

The equipment and set-up were identical to that used in Experiment One.

3.4.2.3 Stimuli

3.4.2.3.1 Audio Stimuli

Twelve different stimuli patterns were used, divided into three groups of four, the groups defined as patterns ending on the same stimulus tone. An audio stimulus pattern consisted of an initial note followed by the actual target note. The two-note phrases are supplied in musical notation in Figure 3.6. The patterns were:

- G_3 target note, preceded by a B_3, F_4, F_3, or C_3;

- D_4 target note, preceded by a Ef_4, D_5, Bf_3, or F_3;
- Cs_5 target note, preceded by a Fs_5, As_5, G_4, or D_4.

Figure 3.6. Two-note phrase stimuli for Experiment Two. The circles appear with the onset of the second note of the phrase. Note that the last four stimuli are all sounded one octave above their notation (indicated by the 8va).

Within each group, the stimulus tone was approached twice from above, and twice from below, known as the descending and ascending phrases, respectively. In addition, each direction of approach was further divided into a large interval and a small interval (greater than or equal to and less than a perfect fifth, respectively). Each of these twelve separate intervals were presented to the subject four times with randomly chosen circles, for a total of 48 stimulus exposures.

3.4.2.3.2 Visual Stimuli

Two grey circles varying in grey level from 0.1 to 0.9 (postscript), each two separated by an interval of 0.4. Luminance of the grey levels in terms of cd/m^2 is supplied in Table 3.1.

3.4.2.4 Procedure

All subjects in Experiment Two had completed Experiment One. Twenty (20) of the subjects had also previously completed Experiment Three. The subjects were told that they would hear first one tone followed by a second tone through the headphones. They were instructed only "...to select the circle that you feel is most like the *second tone that you will hear*. The circles will appear at the same time that you hear the second tone. If you feel that the one on the left is more like the tone, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of tone and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The experiment proceeded with each two note phrase being produced four times, ordered randomly. The program randomly selected a grey level from one of the two circles that ranged in grey value for 0.1 to 0.5. The selection of the one circle determined the grey level of the second circle, as the second circle was offset by a constant difference of 0.4 for discriminability, as described above. The left or right positioning of the two circles was also randomised on each trial. The subject was thus provided with forty-eight separate trials of tone and circle pairs. After the last trial, the screen went blank, and a "Thank You" message appeared. At this point, the subject proceeded either on to Experiment Three (if they had not yet performed the condition) or Experiment Four if they had completed all of the monochromatic experiments (One, Two, and Three).

3.4.3 Results

Figure 3.7 displays the mean scores and standard errors at each stimulus tone level for each condition. These scores are collapsed across subjects, and for all four separate sub-levels for that stimulus tone. Since each target note (TN) was given under the two sub-conditions of direction of approach (ascending, or **A**, and descending, or **D**) and size of leap (small, or **S**, and large, or **L**), each stimulus will be referred to as a combination of its approach characteristics and tone name. Thus, the label **ALG3** refers to the stimulus of a TN of G_3 approached by a large, ascending interval. These labels are used in the tables and figures hereafter.

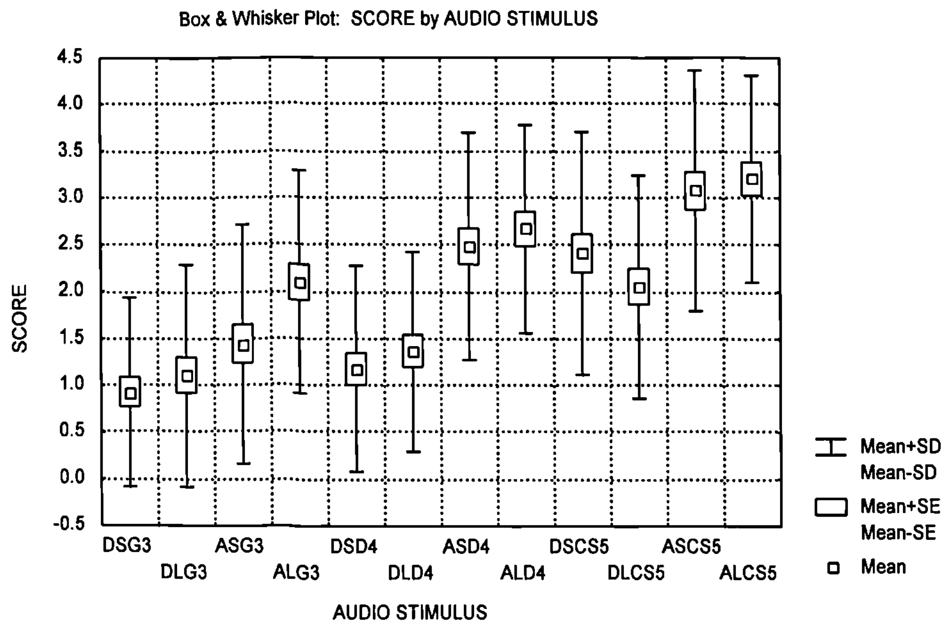


Figure 3.7. Mean scores for each audio stimulus in Experiment Two.

A 2 x 2 x 3 repeated-measures ANOVA was run on the subject brightness scores over the independent variables of direction of intervallic leap (ascending vs. descending), size of leap (large vs. small), and pitch height of the TN. The ANOVA showed a highly significant main effect for TN pitch height ($F(2, 84) = 33.638, p < .001$) and a highly significant main effect for the direction of approach to the TN ($F(1, 42) =$ 29.989,

$p < .001$). No main effect was found for the size of the intervallic leap. Two-way interactions were found between the direction of approach and the TN ($F(2, 84) = 4.789, p < .05$) and between the direction of approach and the size of the formed interval ($F(2, 84) = 3.891, p < .05$).

Figure 3.8 shows the results collapsed across variables to show the main effect of intervallic leap direction. Figure 3.9 is a histogram of subject scores for each of the TN, clearly demonstrating the strong relationship between brightness rating and TN pitch height.

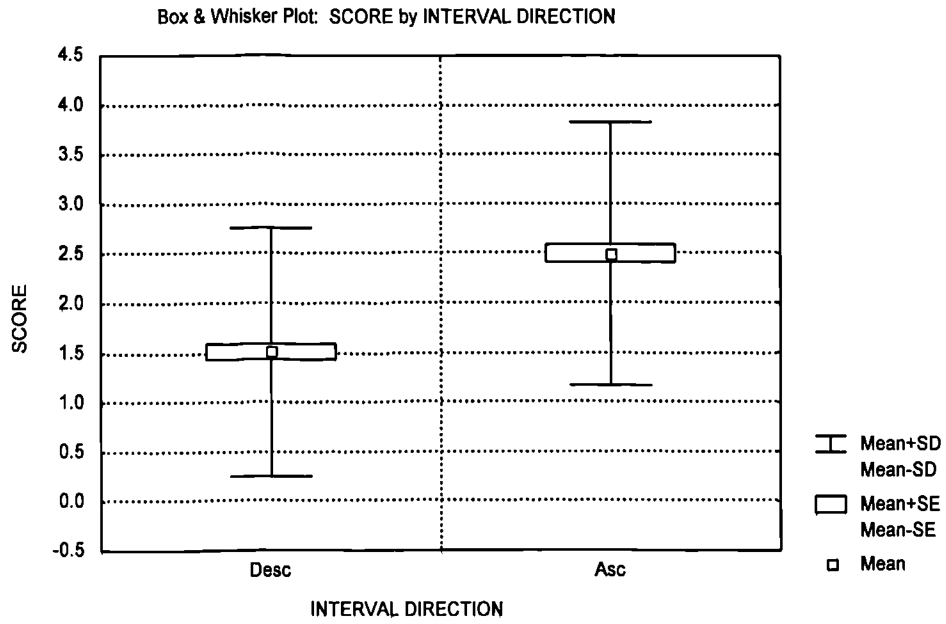


Figure 3.8. Box plot of scores collapsed across variables to show the main effect of interval direction of approach to the TN.

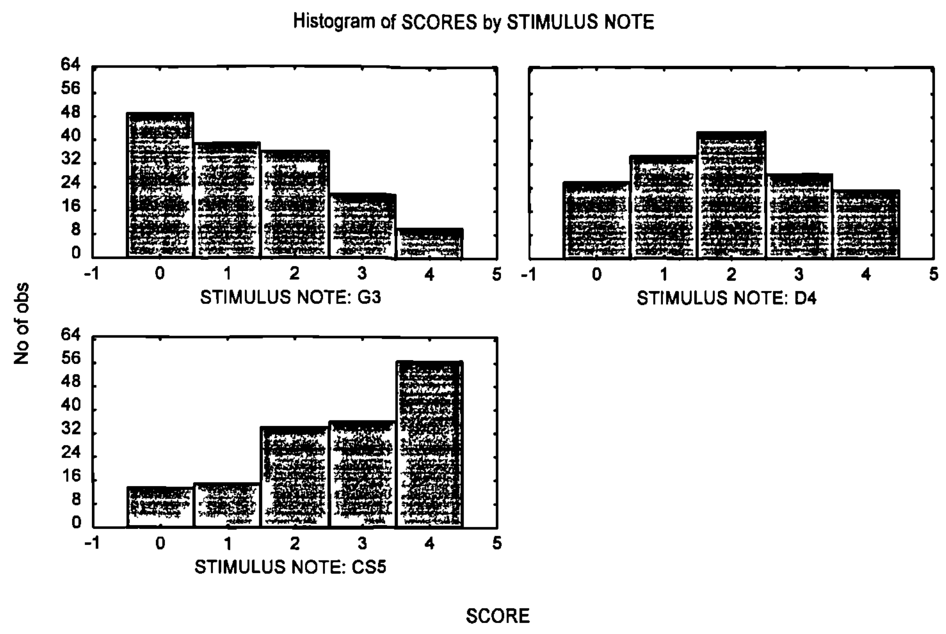


Figure 3.9. Histogram of scores obtained for different stimulus tones. The upper left histogram is for stimulus note G_3; the upper right for stimulus note D_4; the lower left is for stimulus note Cs_5.

For the purposes of this analysis, consonant intervals were specified as thirds, sixths, and perfect intervals, while dissonant intervals were all of the remaining intervals. A histogram of scores obtained for consonant and dissonant interval subsets are shown in Figure 3.10. The histogram reveals that subjects tend to match dissonant intervals with all bright circles or dark circles; consonant interval scores are more normal in distribution. A test of proportions of extreme scores observed (0 or 4) for consonant vs. dissonant intervals reveal a significant difference ($\chi^2(1) = 13.319, p < .001$).

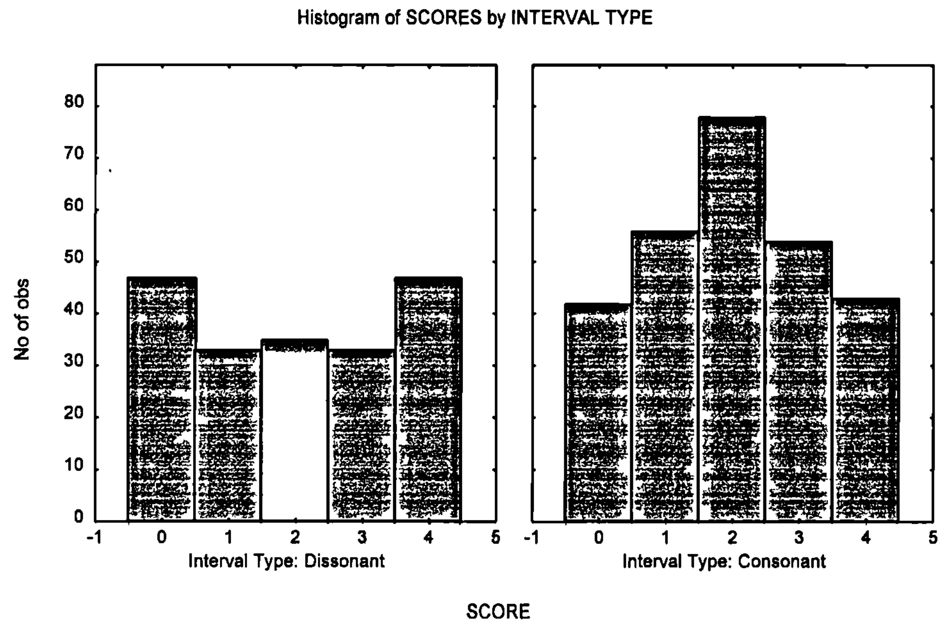


Figure 3.10. Histogram of frequencies of scores obtained for consonant and dissonant intervals.

Individual regressions were run for each subject on their scores and the main factors of the audio stimuli with which they might match circles. These regressions provide information that may not be present in the sample effect. For example, although no main effect is observed for size of intervallic leap, regressions reveal one subject who matched large intervals with brighter circles (Pearson's r = 0.597, $p < .05$), and one who matched small intervals with brighter circles (Pearson's r = -0.784, $p < .01$). Individual subjects demonstrated significant regression coefficients for TN, size of interval, direction of intervallic approach, and TN together with direction of approach. Table 3.3 shows the number of subjects who demonstrated significant regression coefficients for each factor, and the significance level of those coefficients.

p-level	Only Pitch	Only Direction	Only Size	Pitch & Direction
0.05	3/0	4/0	0/1	1/0
0.01	9/0	4/0	1/0	0/0

0.001	0/1	1/0	0/0	0/0
Total	**12/1**	**9/0**	**1/1**	**1/0**

Table 3.3. Number of subjects with significant regression coefficients for each of the main audio factors used to match with circles. Each cell has both the positive and negative regression coefficients, listed as *positive/negative*.

3.4.4 Discussion

The addition of a single preceding tone in this experiment introduces musical content to the stimulus. As observed in Experiment One, the height of the TN had a large effect on the brightness of circle chosen to match the tone. However, this matching was altered by the addition of preceding tone, although subjects were not instructed to use the note in their matching.

The preceding tone serves to alter the rating of the TN according to the direction of the intervallic leap. Thus, leaping upwards to a TN increases the proportion of trials on which the subject matches it with brighter circles, whilst leaping downwards reduces this proportion. This in itself is not too surprising, as no subjects tested had perfect pitch, and relative pitch relationships of notes are more musically important to a listener than the absolute pitch height of those notes. Even in the presence of preceding note context, however, some absolute pitch quality of the tone is not lost; a G_2 would never be considered to be as cross-modally bright as a D_6 in the same manner that a G_2 would never be perceived as high a note as a D_6.

The results of Experiment One suggest that individual subject analysis may reveal cross-modal matching performance that are lost in sample trends. The individual analyses in Experiment Two demonstrate that subjects used different strategies in forming SoCs for the matching task. Although the sample trends reveal that both factors of pitch height and intervallic approach affect cross-modal matching, thirteen subjects use primarily tone height while another nine use only intervallic approach for a SoC. Figure 3.11 illustrates such matching performance by showing regressions lines for one subject's matching and the IVs of TN, consonance, and intervallic approach. The perfect correlation between circle and intervallic approach demonstrates that the subject used only this IV as a SoC.

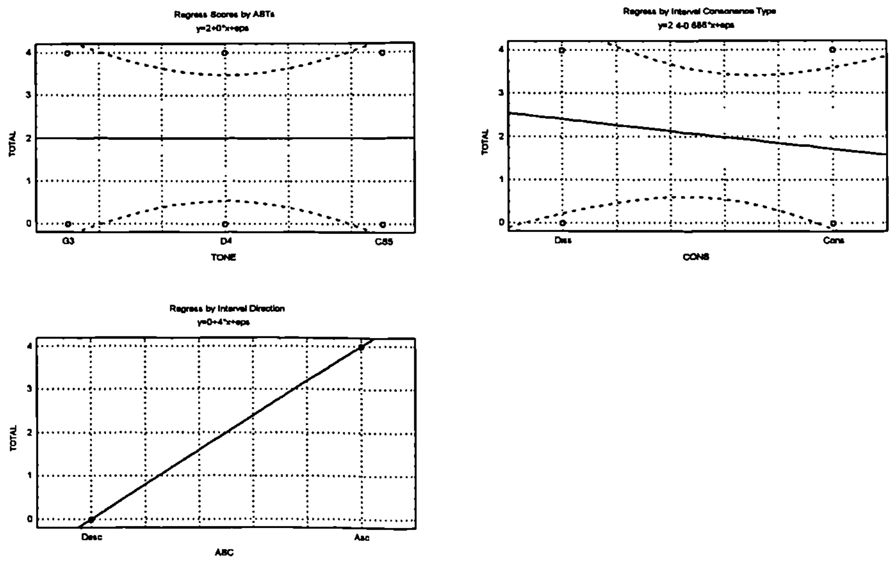


Figure 3.11. Regression of Subject 33's scores and the grouping variables of TN, interval consonance type, and direction of approach to TN.

The simple addition of a preceding tone in Experiment Two creates a number of musical and auditory dimensions. Although not required in the instructions, some subjects use one of these auditory dimensions and translate the single visual dimension of varying grey levels onto it. One subject responds to TN pitch height only, another TN pitch height mediated by direction of approach, and others respond to consonance of formed interval. The direction of alignments between scales may vary between subjects as well.

In addition to revealing strategies that subjects employ to form and use SoCs in Experiment Two, individual analyses may reveal if any subjects have changed matching performance from the first experiment. All but one of the subjects with significant regression coefficients for pitch height and visual brightness had the same direction of matching that they exhibited in Experiment One, including "anomalous" Subject 37. The exception was Subject 32, who changed to the more normal alignment of increasing visual brightness matched with increasing pitch height.

The two subjects who used interval size as a SoC with TN pitch height have correlation coefficients with different signs; that is, they used the same audio SoC but aligned it in different directions with the visual brightness, or one used visual brightness while the other used contrast as a visual SoC.

Examination of these individual subject regressions across independent variables does reveal significant matching between circle brightness and the TN pitch level, intervallic approach direction, and combinations of the two factors. In addition, the independent variable of size, which does not show up as a main effect in the ANOVA, is used as a SoC by at least two subjects. The remainder of subjects are likely influenced by a combination of musical factors in their evaluations.

The results for consonant versus dissonant intervals reveals that subjects are sensitive to musical factors in the experiment in addition to more simple auditory factors already discussed. If only subjects scores for these intervals are examined, an almost identical mean and very small standard error are observed. Yet, a qualitative difference exists in the ratings given to these intervals: subjects are more likely to be consistent in their matching of circles with dissonant intervals than consonant ones. Scores of 0 and 4 are only obtained if the subject chooses the brighter or the darker of the two presented circle at every presentation of the experiment. The more consistent cross-modal matching indicates that the formed interval characteristics of dissonant intervals map more regularly onto the visual SoC of brightness.

The finding of both single and multiple factor SoC matching suggests that the associations of pitch height and brightness might have been an artifact in past experiments. In an effort to control for different factors, tasks were created that yielded results with very strong associations between the limited stimuli. When the stimuli are increased in complexity (and more ecologically valid), a greater variety of response possibilities are created.

3.5 Experiment Three: Cross-modal Matching of Monochromatic Colours and F-chord Variants

3.5.1 Aims and Introduction

Experiment Two further extended cross-modal brightness matching by using musical stimuli even more complex than individual tones or two-note phrases. Another possible enrichment of audio information that may add to the stimuli's musicality is the simultaneous presentation of several constituent tones forming a *chord.*

A chord is poly-dimensional, with underlying dimensions that make it difficult to order linearly. Two chords may differ in only one internal note but have a completely different "character". Using this type of stimuli pushes the SoC paradigm further, and asks if subjects can still form consistent cross-modal matches with audio stimuli varying in many more psychological dimensions.

Experiment Three examined cross-modal matching of chordal stimuli with the same visual stimuli from the earlier experiments. All chordal stimuli were formed from the same fundamental note, in musical theory terms. They were all variants of an F-chord, with all of their constituent tones restricted to one octave (*i.e.*, F_4 to F_5). The chords varied slightly in bottom and top note pitch height. A main effect of pitch height was hypothesised to influence brightness matching performance. However, musical factors specific to chords were also predicted to alter cross-modal matching.

3.5.2 Method

3.5.2.1 Subjects

The subjects consisted of the same 39 individuals who participated in Experiments One and Two.

3.5.2.2 Equipment

The equipment and set-up were identical to that used in Experiment One.

3.5.2.3 Stimuli

3.5.2.3.1 Audio Stimuli

Twelve different stimulus chords were used, all variants of a basic F-chord (although enharmonic analyses of these chords as other chord names are possible). Each of these twelve separate chords was presented to the subject four times with randomly chosen circles, for a total of 48 stimulus exposures. The chords are supplied in a textual format in Table 3.4.

1) F	(F_3, A_3, C_4)	7) Fdim7	(F_3, Af_3, Cf_4, Eff_4)
2) Fm	(F_3, Af_3, C_4)	8) Fhdim7	(F_3, Af_3, Cf_4, Ef_4)
3) F+	(F_3, A_3, Cs_4)	9) F6 3	(A_3, C_4, F_4)
4) Fmaj7	(F_3, A_3, C_4, E_4)	10) Fm6 3	(Af_3, C_4, F_4)
5) Fm7	(F_3, Af_3, C_4, Ef_4)	11) Fm6 5	(Af_3, C_4, Ef_4, F_4)
6) F7	(F_3, A_3, C_4, Ef_4)	12) F6 5	(A_3, C_4, Ef_4, F_4)

Table 3.4. Text description of chordal stimuli used in Experiment Three.

3.5.2.3.2 Visual Stimuli

Two grey circles varying in grey level from 0.1 to 0.9 (postscript), each two separated by an interval of 0.4. The brightness expressed in cd/m^2 is given in Table 3.1.

3.5.2.4 Procedure

Subjects performing Experiment Three had all completed Experiment One, and half of the total subjects had already completed Experiment Two. The subjects were told that they would hear a chord through the headphones. They were instructed only "...to select the circle that you feel is most like the sound. If you feel that the one on the left is more like the sound, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of sound and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The experiment proceeded with each chord being produced four times, ordered randomly. The program randomly selected a grey level for one of the two circles that ranged in grey value for 0.1 to 0.5. The selection of the one circle determined the grey level of the second circle, as the second circle was offset by a constant difference of +0.4 for discriminability, as described above. The left or right positioning of the two circles was also randomised on each trial. The subject was thus provided with forty-eight separate trials of tone and circle pairs. After the last trial, the screen went blank, and a "Thank You" message appeared. At this point, the subject proceeded either on to Experiment Two (if they had not yet performed the condition) or Experiment Four.

3.5.3 Results

Figure 3.12 presents the results for Experiment Three in a box plot of mean scores, standard deviations, and standard errors of scores obtained for each different chord type. A one-way repeated-measures ANOVA examining the scores given to each different chord type reveals a highly significant difference in matching ($F(11, 418) = 14.377$, $p < .001$). The ordering in the graph is by bottom, then top note height. The first eight chords are all in root position, sharing the same base note, and are thus ordered by top note. When two chords share the same bottom and top note, they are ordered with the minor chord first, followed by the major chord.

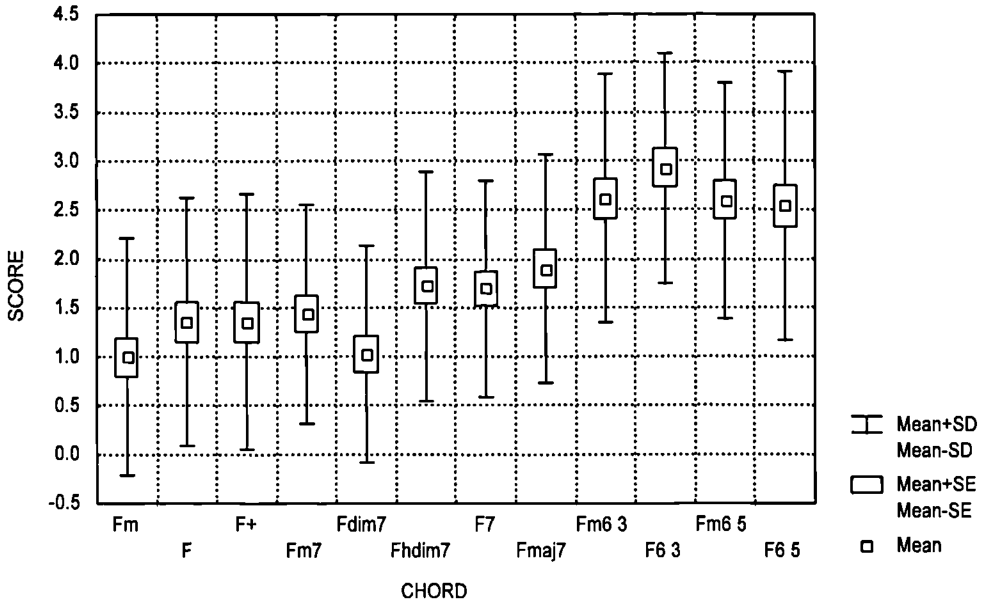


Figure 3.12. Mean scores and standard errors for each chord type. Chords are ordered by increasing bottom note pitch height, then increasing top note pitch height.

Planned comparisons run on the ANOVA results reveal significant differences between root position chord and first inversion chords ($F(1, 38) = 38.241, p < .001$). Comparisons between scores recorded for individual stimuli reveal significant differences between inverted chords and their root position counterpart. F6 5 was brighter than F7 ($t(76) = 3.00, p < .01$), Fm6 5 was brighter than Fm ($t(76) = 5.80, p < .001$), F6 3 was brighter than F ($t(76) = 5.66, p < .001$), and Fm6 3 was brighter than Fm ($t(76) = 5.74, p < .001$). Comparison of chord tonality was performed by grouping F and Fmaj7 as major chords, and testing them against Fm and Fm7 as minor chords. This analysis found that major chords are rated brighter than minor chords ($t(154) = 2.12, p < .05$).

No differences are observed between three- and four-note chords. Figure 3.13 shows the individual histograms of subject scores (number of times the lighter circle chosen) for each of the twelve chord types.

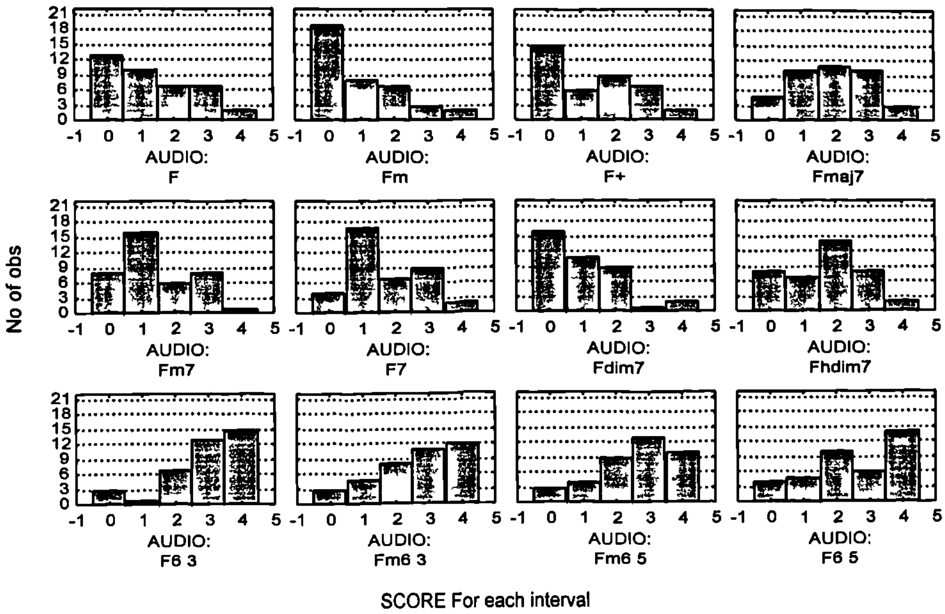


Figure 3.13. Histograms of scores for each chord type. The abscissa for each histogram shows the score recorded for the chord.

Individual regressions were run for each subject on their scores and the main factors of top note height, bottom note height, and number of tones in the chord. These regressions provide information that may not be present in the sample effect. Although it is difficult to divide chords into consonance groups, a limited subset of the chords can be divided into groups: major and minor chords in root position and first inversion (Stimuli 1, 2, 9, & 10). Additional regressions were run the major and minor chord subsets. Table 3.5 shows the number of significant regression coefficients for each of the IVs and the sign of those coefficients. Of the total 39 subjects, 25 subjects (69%) demonstrate at least one significant regression between their scores and an IV.

p-level	Only *Bottom Note*	Only *# of Notes*	*Consonance*	*Top* Over *Bottom*	*Bottom* Over *Top*	*Total* Over *# of Notes*
0.05	4/0	0/1	0/0	2/0	1/0	0/0
0.01	2/0	0/0	1/0	3/0	4/0	0/1
0.001	0/0	0/1	0/0	2/0	3/1	0/0
Total	**6/0**	**0/2**	**1/0**	**7/0**	**8/1**	**0/1**

Table 3.5. Number of subjects with significant regression coefficients for each of the main audio factors used to match

with circles. Each cell has both the positive and negative regression coefficients, listed as *positive/negative*.

3.5.4 Discussion

The stimuli are more complex than those in the previous two experiments; the analyses still reveal that subjects are making regular cross-modal matchings, but now availing themselves of the additional audio stimulus dimensions in their SoCs. As in Experiment Two, individuals use either one primary auditory factor or a combination of factors to form their matchings. In at least one case, subjects are found that use the same scales, but are aligned in opposite directions.

The majority of subjects that have significant individual regression coefficients use a combination of top and bottom note height and aligned these factors so that increasing pitch height is matched with increasing visual brightness. The same predominant alignment is thus exhibited in all three experiments. The same subject that showed opposite alignment of brightness or use of contrast as her SoC in the first two experiments continues to demonstrate the same matching pattern.

Regressions indicate that some subjects primarily use either of the single factors of top note height or bottom note height to match with the visual stimuli. The chords vary in top note height by only a perfect fifth, but subjects still exhibit extreme scores (0 and 4) for the extremes of the audio stimuli. That is, they demonstrated a full range of variation in responses as they did in the other experiments. This again supports the hypothesis that matchings are formed by using SoCs formed from ranges of presented stimuli scales rather than an absolute relationship between visual brightness and pitch height.

The highly significant main effect found for root position vs. first inversion is most likely due to the absence of bottom frequencies. When scores between the same chord in root position and inversion are compared, each pair reveals significant differences. For one pair, these chords differ in top note height by the smallest possible musical interval: a minor second. This indicates that the absence of those lower frequencies (and presence of higher frequencies from the inverted F) is a large enough change in quality to alter cross-modal matching.

Some subjects used an audio factor that may be called chordal complexity, made up of the number of constituent notes in the chords. For all subjects that used complexity for the matching, the chords with fewer notes were matched with higher visual brightness.

3.6 Conclusions

The use of a two-alternative forced choice as an experimental procedure does reveal agreements in standard cross-modal matching experiments. In accordance with earlier studies, a strong tendency to match brighter colours with higher tones was observed. This agreement was demonstrated with the use of a computer VDU for visual stimuli presentation, and the use of complex waveform musical tones as auditory stimuli.

Cross-modal matching performance could first be analysed by ANOVA, which reveals significant effects of differences in the stimulus conditions across the sample. Taken as a sample, systematic matchings between acoustic dimensions and visual brightness can be demonstrated.

Individual subject performance was then analyzed to establish the strategies used by individual subjects to carry out the matching. Different SoCs were used by different subjects; in some cases several subjects used the same SoC, but aligned it to the source scale in opposite directions.

In Experiment One, most subjects demonstrate regular matching of higher pitch height tones with brighter circles, replicating previous experiments. In the first experiment, with the simplest, least ambiguous acoustic stimulus, two subjects responded in the opposite direction. A single suprasensory representation of brightness cannot account for the matching performances of these subjects.

Experiment Two examined cross-modal matching performance when the audio stimuli were increased in complexity by adding a preceding tone. A main effect for the pitch height of the stimulus note remained, but this effect was modulated significantly by the preceding tone in various ways. Although this preceding tone was not to be judged, the results demonstrated that subjects did not ignore it. The direction that target note was approached from affected the brightness of the circle with which it was matched. The implication is that the perceived pitch height, or even perceived acoustic brightness of the target note is adjusted away from the direction that it is approached. Analysis of sample trends also revealed that subjects were more likely to match formed dissonant intervals with all brighter circles or all darker circles.

Analysis of individual subject matching performance in Experiment Two revealed that different subjects used different single factors for matching. One of those factors, formed interval size, is revealed in individual performance analysis, but lost in the sample trend. As the number of possible dimensions in the auditory stimuli that could be used increased, the number of SoCs that subjects did use also increased. If there were a single suprasensory dimension that subjects use, this would not be so.

In Experiment Three even more complex auditory stimuli were used. The number of possible factors that subjects could use in SoC creation is much larger. Regular matching performance for the sample is still shown for pitch height and brightness of circle matched. Analysis of individual subject performance revealed that some subjects again used one or several auditory stimulus dimensions to form cross-modal matches.

The three experiments demonstrate that subjects continue to demonstrate cross-modal matching in the face of more complex auditory stimuli; however, the strategies that are used by the subjects increase along with increasing stimulus dimensions. As the auditory stimuli increase in complexity, they become less musically ambiguous, and should be more familiar to subjects. Even so, variance in cross-modal matching performance increases rather than becoming uniform across the sample.

What important qualities do dimensions have that make subjects use some and not others? Most likely, it is the accessibility of the dimension that predisposes a subject to use it as part of a SoC. It is much easier to access the qualities of relative pitch height of a note or the intervallic direction of approach than the size of an interval or more complex musical dimensions of consonance or chord type. If a subject were more accustomed to using a dimension regularly in life, she would have that dimension available as a SoC for cross-modal comparisons. Thus, a highly trained musician might use a scale of consonance quite regularly to form matches.

4 Cross-Modal Matching of Simple Musical Constructs and Coloured Visual Stimuli

4.1 Overview

This chapter extends the results of the grey-level cross-modal experiments by varying the hues of the visual stimuli presented to subjects. The impetus for such an extension arises from the lack of systematic examination of hue in such matchings, and the strong presence of a hue component in synaesthetic imagery. Most research to date has been conducted with visual stimuli varying in grey-levels only. These experiments also examined the role that affective response may play in mediating cross-modal matching. Regular matching between a colour's luminance and a single tone's pitch height was observed when colours were roughly matched for brightness and saturation. Regular matching across subjects broke down with two-note intervals, although some regular patterns between musical characteristics of the interval and brightness of colour was observed with individual subjects. The matching of chords and colours revealed similar effects to the grey-level matchings; primarily, that inverted chords were routinely matched with brighter circles. These results demonstrated that cross-modal matching was still possible with more complex coloured visual stimuli, although the consistency of such matchings was reduced.

4.2 Introduction

The previous chapter reported that matching the brightness of visual stimuli to musical constructs could be used as a method of "eavesdropping" on some of the perceiver's representations of the music. When grey circles were presented with tones varying only in pitch height, subjects normally align the SoCs of visual brightness and tone height such that rising pitch aligns with increasing brightness. This alignment was found to be inverted for a small number of individuals. When the musical stimuli increase in complexity, a number of integral dimensions to the music become available for alignment with brightness. These dimensions may singularly, or in combination comprise the SoC that is used to align with the visual SoC.

Synaesthetic imagery almost always contain a hue component along with other specific visual attributes. The results from the previous chapter were extended in the following experiments by varying the hue of the visual stimuli. One motivating factor for this is that specific colour names are among the primary adjectives used by synaesthetes to describe their photisms. This also provided the opportunity for multiple individual scalings and orderings of the stimuli to be examined. Preference for hue may also correlate with affective components. Altering the simple matching experiment by only varying hue provides the initial opportunity to examine the role of affect and hue preference in cross-modal matching. The experiments in this chapter use the same audio stimuli and presentation method to examine the effect of adding the dimension of varying hue, roughly balanced for brightness and saturation, to the matching task. Personal ratings of these colours were gathered from each subject to allow for analysis of matching on both objective and subjective scales.

4.3 Experiment Four: Cross-modal Matching of Colours and Single Musical Tones Matching

4.3.1 Aims and Introduction

Experiment Four was identical in nature to Experiment One, with the addition of varying the hues of the stimulus circles. All subjects from the Chapter Three experiments proceeded to it after completing the three monochromatic conditions. The primary intention of the experiment was to discover if the pitch height and visual brightness alignment, observed in Experiment One, would still be present when the circles were coloured with varying spectral hues. Experiments Two and Three demonstrated that cross-modal matching can change when the auditory stimuli become more complex. This experiment performed the same purpose in examining how SoCs change in different tasks when the visual stimuli are made more complex.

The term *colour* has been routinely applied to visual stimuli used in cross-modal matching experiments for decades. In cross-modal literature, *colour* has most often referred to white, black, and grey visual stimuli (Ellson, 1941a & 1941b; Marks, 1966, 1974; Melara, 1989a, 1989b). Some experiments have attempted to examine the contribution of hue to cross-modal judgments (Wicker, 1968; Simpson, Quinn, & Ausubel, 1956). Almost all of these experiments examined hue independent of other colour attributes, such as brightness and saturation. This pre-supposes that individuals across the population will have identical or similar perceptual experiences in response to specific hues. A notable exception is a cross-modal study conducted by Rader & Tellegen (1987) which examined cross-modal matching of tones to colours. The colour responses were verbally provided by subjects without reference to an actual physical colour stimulus. For analysis, the experimenters ordered the colours by their brightnesses, but did not provide an indication of how those brightness levels were decided (*i.e., in order: white, pink, yellow, red, orange, green, blue,* purple, brown, grey, black). Analyzed on this scale, low tones (200 Hz) were matched with dark colours, medium tones (1000 Hz) were matched with middle brightness colours, and high tones (4000 Hz) were matched with the bright colours. Using this ordering of colours, the colour matching that young children performed in the Simpson, Quinn, & Ausubel experiment (1956) follows the same pattern, with children matching low tones with blue and violet, medium tones with orange and red, and high tones with yellow and green.

This experiment introduced several varying hues in a balanced manner to determine if hue itself may serve as an deffective SoC. If it does not, it may be possible then to determine which other attributes of the visual stimuli were employed as a SoC, or whether introducing varying hues greatly changes the nature of cross-modal matching performance.

The use of hue as well as brightness and saturation is necessary in order to make comparisons between cross-modal matching and synaesthetic experiences. Synaesthetes anecdotally report that higher tones are brighter in colour (Cytowic, 1989), but also frequently report hues as a fundamental component of their imagery. Visual synaesthetes themselves have idiosyncratic correspondences between colour imagery and invoking stimuli. The individual nature of these correspondences suggest that hue will not be used as an SoC in an identical manner for all non-synaesthetic subjects.

4.3.2 Method

4.3.2.1 Subjects

The subjects consisted of all of the participants of Experiments One, Two, and Three.

4.3.2.2 Equipment

The equipment and set-up were identical to that used in Experiment One.

4.3.2.3 Stimuli

4.3.2.3.1 Audio Stimuli

The twelve Western tonal notes, covering a range of over three octaves, ordered as: F_3, A_3, Cs_4, E_4, G_4, Bf_4, C_5, Ef_5, Af_5, B_5, D_6, Fs_6. This set of auditory stimuli is identical to the set from Experiment One, and is shown in musical notation in Figure 3.1 (p. 3–4). Each stimulus was presented with randomly chosen circles four times.

4.3.2.3.2 Visual Stimuli

Two coloured circles were presented side by side in each trial, with each circle subtending a 13° visual angle. The visual stimuli were prepared on the NeXT display using the HSB model of colour description. The exact levels of each of these colours is given in Table 4.1. The colours were divided into three sets; one set at full brightness and saturation, one at 50% brightness and full saturation, and one at 50% saturation and full brightness. Each set was composed of six colours, each composed of one of the primary and secondary hues (red, orange, yellow, blue, green, violet). Thus, each of these six hues was presented at three saturation/brightness levels, comprising a total set of 18 distinct colours. Because of the great number of pairings that would be possible with 18 colours (152), and to control the presentations for controlled judgments made by varying hue, each colour was presented with another colour from its same saturation/brightness group.

	Colour	Candelas per square metre	CIE X	CIE Y
R	Red	19.8	0.536	0.390
E	Orange	26.7	0.490	0.431
G	Yellow	55.2	0.409	0.502
	Green	46.9	0.350	0.549
	Blue	15.6	0.203	0.139
	Violet	18.7	0.233	0.155
L	Red	37.7	0.364	0.339
I	Orange	42.1	0.355	0.358
G	Yellow	58.6	0.350	0.400
H	Green	49.1	0.315	0.406
T	Blue	53.0	0.319	0.407
	Violet	32.1	0.246	0.226
D	Red	11.4	0.476	0.429
A	Orange	12.3	0.466	0.440
R	Yellow	17.7	0.424	0.485
K	Green	17.3	0.403	0.497
	Blue	11.1	0.331	0.292
	Violet	10.7	0.324	0.291

Table 4.1. Candelas per square metre and (X, Y) coordinates in CIE colour space for each of the eighteen colours used for Experiments Four–Six.

4.3.2.4 Procedure

Subjects performing Experiment Four had completed all of the monochromatic matching experiments. The subjects were told that they would hear a tone through the headphones. They were instructed only "...to select the circle that you feel is most like the tone. If you feel that the one on the left is more like the tone, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of tone and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The experiment proceeded with each tone being produced five times, ordered randomly. The program randomly selected a colour for one of the two circles from the available colours at a random saturation/brightness level. The selection of the second circle was randomly chosen from amongst the remaining five colours at the same saturation/brightness level as the first colour. The left or right positioning of the two circles was randomised on each trial. The subject was thus provided with sixty separate trials of tone and circle pairs. After the last trial the screen went blank, and a "Thank You" message appeared. At this point, the subject proceeded either on to Experiment Five or Experiment Six, selected in a random fashion.

At the completion of all three hue-based cross-modal experiments (Four, Five, and Six), the subjects each performed a rating task, individually rating each of the visual stimuli used in the experiments. Since these ratings were crucial to the scoring of individual results for all three experiments, the procedure for collecting these ratings will be discussed here, out of chronological order.

4.3.2.5 Ratings

At the completion of all three experiments, the subjects were requested to rate each auditory and visual stimulus on two linear scales. The presented scales were *Bright/Dim* and *Attractive/Not Attractive.*

For each visual stimulus, the subject saw the circle in the centre of the screen and was asked to rate the colour on the two scales of *Attractive/Not Attractive* and *Bright/Dim*. The subject was presented with a horizontal slider that they could position by using a mouse to supply their response. This slider stored a floating number ranging from -1 (*Not Attractive* or *Dim*) to 1 (*Attractive* or *Bright*), repositioned at 0 (*Neutral*) for the beginning of each rating trial (see Figure 4.1). After making the two ratings, the subject pressed a button entitled "Rate Colour", which then presented the next colour. If the subject wished to give a colour a neutral rating on either of the two scales, they were first required to move the scale slider and return it to a neutral position, or they would not be permitted to proceed to the next stimulus. This restriction prevented subjects from merely stepping through the rating task as an exercise.

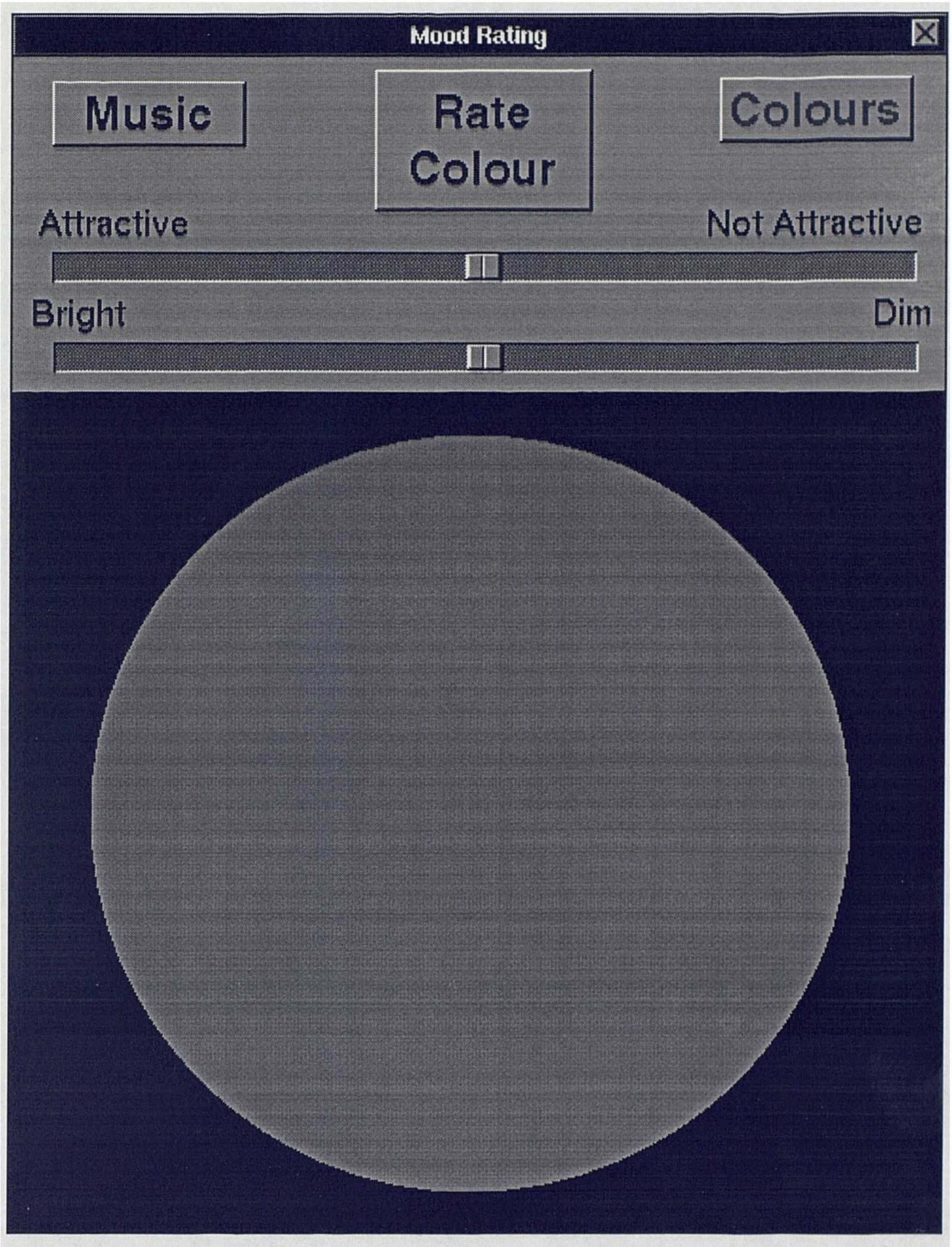


Figure 4.1. Colour mood rating window. Colour ratings are provided by the subject moving each of the sliders on the top of window, and pressing the *Rate Colour* button.

4.3.3 Results

Scores were computed for each audio stimulus, using each of the three previously described scales. The value for the colour along one of these scales was substituted for that colour name in the data to compute the score. Thus, for the subjective brightness ratings, a number from 0 to 5 can be computed for each subject's matching of an audio stimulus. A score of 0 represents that the subject always selected the subjectively darker circle on each presentation of the tone; a score of 5 represents the brighter of the two circles always being selected. In this same fashion, scores were computed for the selection of the more attractive circle and the selection of the more luminous circle.

The cross-modal matching could also be analysed on a scale of colour order, although it is difficult to hypothesise *a priori* what type of ordering different subjects might use. Since no such scale exists, the colours were ordered on the physical scale of their spectral frequencies: *red, orange, yellow, green, blue,* and *violet* (ROYGBV). This order was substituted for the colour names in the data, and a spectral ordering score was computed in the same fashion as the luminance scores.

Figure 4.2 shows the mean scores and standard errors computed from individual subject subjective brightness ratings for each audio stimulus ordered by pitch height. Statistics were performed on scores computed from each individual subject's subjective ratings. A one-way repeated measures ANOVA run on these scores reveals a highly significant effect ($F(11, 418) = 3.622; p < .001$).

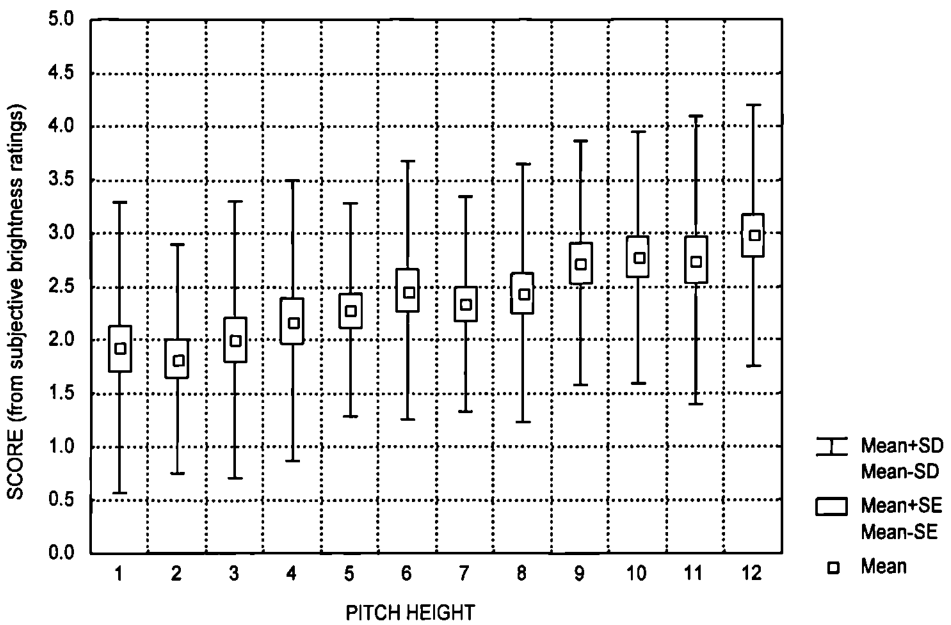


Figure 4.2. Mean scores for subjective brightness rated colour matching with pitch height.

Examining the same matching performance using luminance as the discrimination scale for colours reveals the same trend, but more pronounced. Figure 4.3 plots the mean scores against pitch height when scores are computed from luminance values. A correlation of luminance values and pitch height also reveals a highly significant linear relationship (Pearson's ρ = 0.524, $p < .001$). A one-way repeated measures ANOVA of luminance values by audio height yields a higher F-ratio, also highly significant

($F(11, 418) = 17.895; p < .001$).

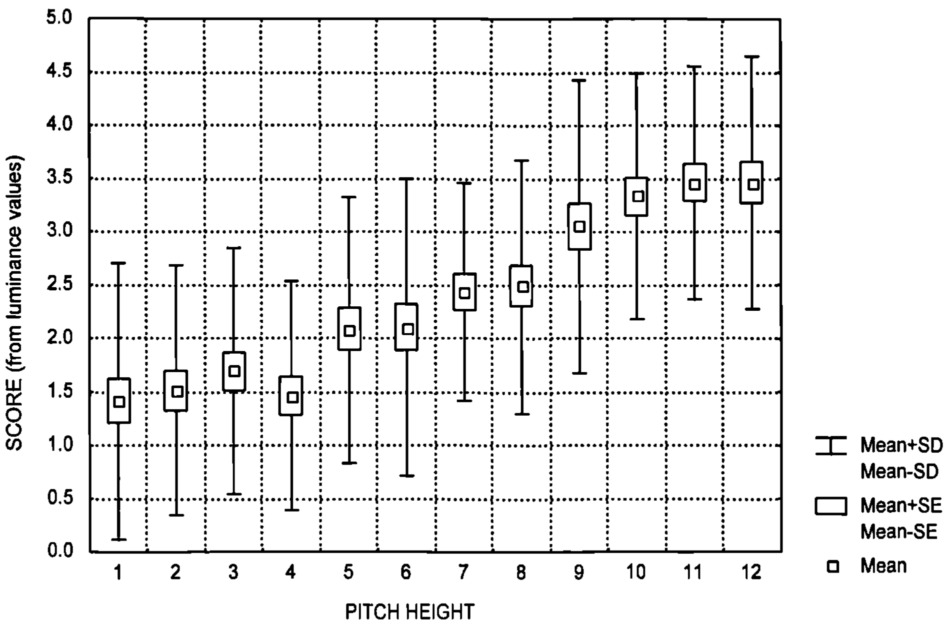


Figure 4.3. Mean scores for colour circle / pitch height matching when circles are rated by luminance, measured in candelas per square metre.

The scores can also be computed from the attractiveness ratings that subjects provided at the completion of the three colour cross-modal matching experiments. Figure 4.4 plots the means scores of the more attractive circle selected in the ordinate by pitch height in the abscissa.

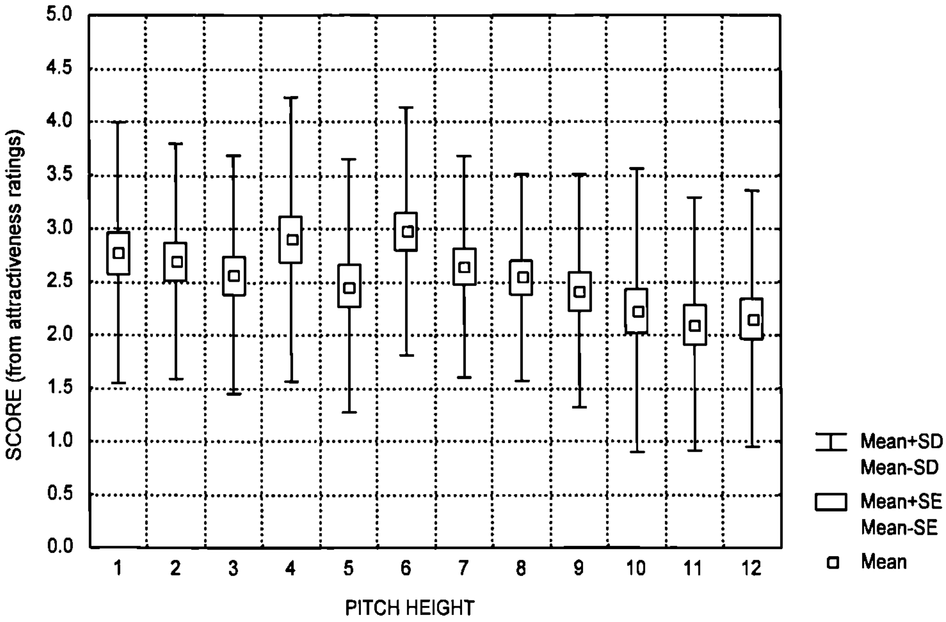


Figure 4.4. Mean scores for colour circle / pitch height matching when circles are rated by individual subject attractiveness ratings.

The analysis by spectral order (ROYGBV) does not yield any main effects. In order to determine inter-subject reliability in the subjective brightness ratings of the stimuli colours, Kendell's Coefficient of Concordance was computed and demonstrates significant agreement between the different subjects ($W = 0.423$; $\chi^2 = 280.250$; $p < .001$). Significant agreement also exists using subject ratings for attractiveness ($W = 0.256$; $\chi^2 = 169.657$; $p < .001$).

Individual regressions were run for each subject on their scores computed from luminance, spectrum position, brightness ratings, and attractiveness ratings, producing four separate regression coefficients. Table 4.2 shows the number of significant regression coefficients found, with the positive and negative coefficients in the same cell (positive on left, negative on right). The rows in the table shows the number of coefficients found below each significance level. Since two or more scales may be correlated for any subject, the same subject may appear in more than on column.

p-level	Luminance	Spectral	Brightness	Attractiveness
0.05	11/1	1/6	6/2	2/4
0.01	7/1	0/0	2/0	1/2
0.001	7/0	0/0	3/0	0/0
Total	**25/2**	**1/6**	**11/2**	**3/6**

Table 4.2. Number of subjects with significant regression coefficients computed from colour rating scale and circle chosen. Each cell has both the positive and negative regression coefficients, listed as *positive/negative*.

4.3.4 Discussion

With each of the three scoring methods yielding different results, it remains to be seen which of the methods best describes the subject matching. The subjective brightness ratings should agree with the luminance scale if the subject's scale was perfectly correlated with the physical scale. The subjective ratings do, in fact, significantly correlate with the luminance values (Pearson's $\rho = 0.411, p < .001$) for 24 of the subjects (62%, Pearson's $\rho > 0.466, p < .05$). All of the correlations were positive in sign. The question remains as to why all subjects did not demonstrate significant correlations if the scales were similar. Possibly the rating task was too brief in nature, with each of the colours being rated only once. The relationship between increasing pitch height and increasing luminance is stronger than the similar correspondence between pitch height and subjective brightness. This suggests that luminance may provide a more useful scale for examining cross-modal matching than subjective brightness scale in the succeeding experiments.

Attractiveness judgments negatively correlate with the luminance values (Pearson's $\rho = -0.234, p < .001$), and this correlation and the previous one are significantly different ($Z = 12.623, p < .001$). The attractiveness ratings would suffer from the same limitations as the subjective brightness ratings. And, the attractiveness ratings do not account for the pattern of cross-modal matching as well as the luminance values.

Individual analyses again reveal that subjects use different SoCs for forming cross-modal matches. Subjects are found that use each of the three visual SoCs. Most subjects who have significant regression coefficients show matching between luminance and pitch height. A majority of these subjects align increasing luminance with pitch height, but again, as in Experiment One, some do align these scales in the opposite direction.

In Experiment One, the visual stimuli of decreasing brightness aligned with tones of increasing pitch height could also be interpreted as if subjects were matching circles of increasing contrast with increasing pitch height. In Experiment Four this is not the case, as the less luminous circles also have less contrast with the black screen background. In this experiment, those subjects that demonstrate this type of matching performance actually are matching circles of decreasing brightness with increasing pitch height. This provides more evidence that the hypothesised relationship between increasing brightness and increasing pitch height is not universal.

The individual regression using the spectral order of colours did reveal some significant regression coefficients, albeit not as strongly as the other scales. This provides some evidence that some subjects may use a more abstract colour ordering to make these cross-modal matches.

The regression coefficients for scores computed from subjective brightness and attractiveness ratings demonstrate that different SoCs can be used in different alignments to form cross-modal matches when stimulus complexity increases. Although the separate scales for the visual stimuli may be significantly correlated, some subjects have significant regression coefficients for only one of the scales, indicating that the subject was using some varying quality of the visual stimuli as a SoC that is also correlated with one of the experimenter presented scales. Experiments Two and Three in the previous chapter demonstrated that an increase in auditory stimulus complexity results in subjects using more of the qualities in forming cross-modal matches; Experiment Four demonstrates that different visual qualities may be used as SoCs when they become available in the presented stimuli.

4.4 Experiment Five: Cross-modal Matching of Colours and Two-Note Phrases

4.4.1 Aims and Introduction

Experiment Five builds upon the findings of Experiment Two in the same manner that Experiment Four extended the results of Experiment One. This experiment set out to examine how the cross-modal matching of two-note temporal phrases changes when colours of varying hue and saturation are used instead of monochromatic circles. The same scales used for analyzing the colours in Experiment Four were again used here.

4.4.2 Method

4.4.2.1 Subjects

The subjects consisted of all of the participants of Experiments One, Two, Three, and Four.

4.4.2.2 Equipment

The equipment and set-up were identical to that used in Experiment Four.

4.4.2.3 Stimuli

4.4.2.3.1 Tones

Auditory stimuli were identical in nature to those used in Experiment Two.

4.4.2.3.2 Visual

Visual stimuli were identical in nature to those used in Experiment Four.

4.4.2.4 Procedure

Procedure was identical to Experiment Two, which examined grey-circle matching with two-note phrases. The visual stimuli and procedure were identical to Experiment Four.

4.4.3 *Results*

A 2 x 2 x 3 repeated measures ANOVA performed on luminance values by the independent factors of interval direction, size, and height of TN yields no main effects or interactions. No main effects nor interactions are observed in 2 x 2 x 3 ANOVAs performed on the attractiveness, brightness, or spectral position scores by the independent factors of interval direction, size, and height of TN. Individual regressions were again run for each subject on their scores computed from the four colour scales with each of the independent factors of the audio stimuli, producing separate regression coefficients. Table 4.3 shows the number of significant regression coefficients found, with the positive and negative coefficients in the same cell (positive on left, negative on right). The rows in the table show the number of coefficients found below each significance level. Since two or more scales may be correlated for any subject, the same subject may appear in more than one column.

	p-level	Luminance	Spectral	Brightness	Attractiveness
Tone Height (TN)	*0.05*	1/2	0/1	0/2	0/1
	0.01	0/0	1/0	0/0	0/0
	0.001	0/0	0/0	0/0	0/0
	Total	**1/2**	**1/1**	**0/2**	**0/1**
Phrase Direction	*0.05*	1/1	1/0	0/0	2/0
	0.01	0/0	0/0	0/0	1/0
	0.001	0/0	0/0	0/0	0/0
	Total	**1/1**	**1/0**	**0/0**	**3/0**
Interval Size	*0.05*	0/2	0/1	0/2	1/2
	0.01	0/0	0/0	0/0	1/0
	0.001	0/0	0/0	0/0	0/0
	Total	**0/2**	**0/1**	**0/2**	**2/2**
Consonance	*0.05*	1/2	0/1	0/2	0/1
	0.01	0/0	1/0	0/0	0/0
	0.001	0/0	0/0	0/0	0/0
	Total	**1/2**	**1/1**	**0/2**	**0/1**
Grand Totals		**3/7**	**3/3**	**0/6**	**5/4**

Table 4.3. Number of subjects with significant regression coefficients in Experiment Five for each of the main audio factors used to match with circles. Each cell has both the positive and negative regression coefficients, listed as *positive/negative*.

4.4.4 Discussion

Far fewer regular associations between luminance and stimulus tone TN and intervallic direction were found compared to those in Experiment Two. The lack of sample main effects and the diversity in the significant regression coefficients observed for individual subjects indicates that the diversity in subject responses did grow as the visual and auditory stimuli grew in complexity. The visual stimuli in Experiments One, Two, and Three varied along one dimension only, whether that dimension be luminance or contrast. The change from monochromatic grey colours to a larger range of colours creates the opportunity for the hues to be aligned using a number of different SoCs appropriate for that individual. In order to determine what SoCs were being used in the matchings, it would be necessary to either determine how the subject was ordering the stimuli, or provide more scales for response. The latter technique is utilised in Chapter Six, in order to examine matchings between higher order musical passages and affective meaning.

The small number of significant regression coefficients in Table 4.3 for each visual and auditory factor makes it impossible to assert which factors are the most useful in cross-modal matching. The results of the experiment suggest that if both sets of stimuli presented in a cross-modal experiment vary too broadly in their dimensionality, the individual matching performances become greatly varied and idiosyncratic. Such varied matching is difficult to observe in experiments, and may be part of the reason that simple stimuli have been used in past cross-modal experiments.

4.5 Experiment Six: Cross-modal Matching of Colours and F-chord Variants

4.5.1 Aims and Introduction

Experiment Six builds upon the findings of Experiment Three in the same manner that Experiment Four extended the results of Experiment One. In light of Experiment Five's results, a doubt arises as to whether increasing the complexity of audio stimuli from a single note to a chord will still yield significant cross-modal matching. The increase in complexity of the visual stimuli in Experiment Five greatly reduced the number of observable cross-modal matching performances when compared to Experiment Two. This experiment will determine if a similar decrease in consistent matching performances occurs when compared to Experiment Three.

The chords used in this experiment are more musically complex than the individual tones from Experiment Four. However, they do not have the temporal qualities of the intervals in Experiment Two and Five. Those temporal intervals had many potential qualities that might be used in creation of a SoC. The more static nature of isolated chords may have quite different qualities that operate as more effective SoCs. This potentially allows for the extension of results from Experiments One and Three despite the difficulties observed with the temporal intervals.

4.5.2 Method

4.5.2.1 Subjects

The subjects consisted of all of the participants of Experiments One, Two, Three, Four, and Five.

4.5.2.2 Equipment

The equipment and set-up were identical to that used in Experiment Four.

4.5.2.3 Stimuli

4.5.2.3.1 Tones

Auditory stimuli were identical in nature to those used in Experiment Three.

4.5.2.3.2 Visual

Visual stimuli were identical in nature to those used in Experiment Four.

4.5.2.4 Procedure

Procedure was identical to Experiment Three, which examined grey-circle matching with F-chord variants. This experiment differed in that the circles were coloured. The visual stimuli and procedure were identical to Experiment Three.

4.5.3 Results

A one-way repeated measures ANOVA reveals significant differences between the scores for each of the chords, computed from the luminance values for the circles ($F(11, 418) = 10.346$, $p < .001$). Figure 4.5 shows the mean scores computed from luminance values for each colour stimulus. A planned comparison on the means reveals a significant difference between brightness ratings for chords in root position and first-inversion ($F(1, 38) = 33.727$, $p < .001$). The same one-way repeated measures ANOVA was run on the results using the attractiveness rankings for computing scores, and reveals a significant difference in rankings with a smaller F-ratio
($F(11, 418) = 2.321$, $p < .01$). The mean scores computed from attractiveness rankings is plotted in Figure 4.6. Planned comparisons for the means between groups do not reveal a significant difference between ratings for root position and first-inversion chords, but do reveal a significant difference between ratings for three and four note chords ($F(1, 38) = 173.099$, $p < .001$). No main effects are observed for the scores computed from the spectral ordering scores.

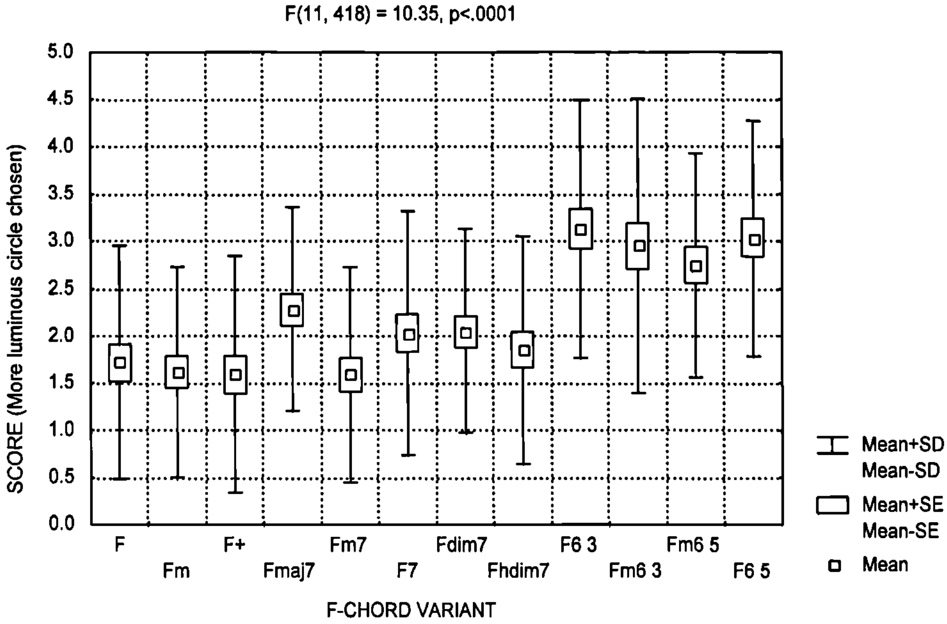


Figure 4.5. Mean scores for cross-modal matching of F-chord variants and circles. Scores are computed from luminance values for each colour.

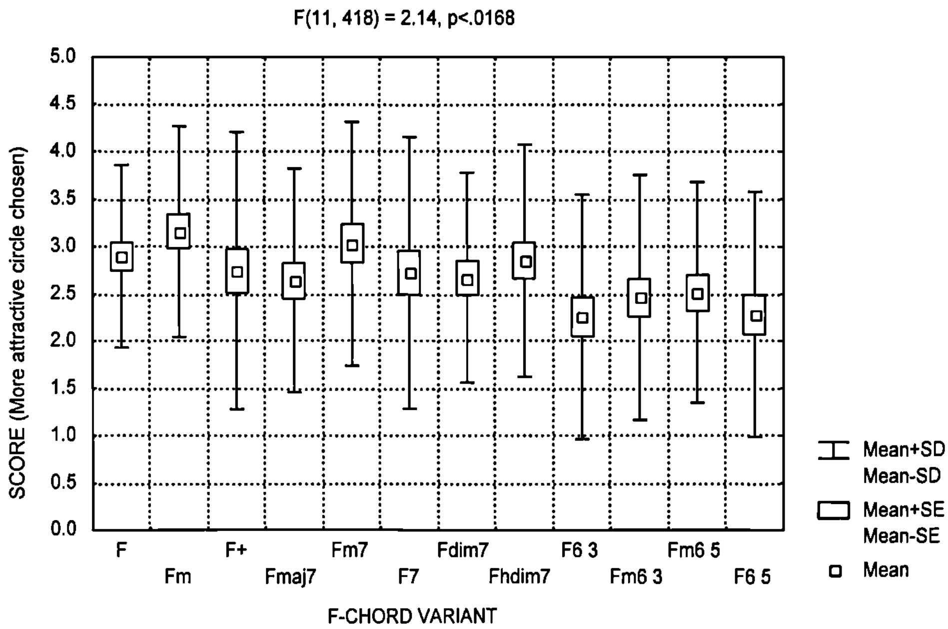


Figure 4.6. Mean scores for cross-modal matching of F-chord variants and circles. Scores are computed from attractiveness ratings provided by each subject for each colour.

Individual regressions were again run for each subject on their scores computed from the four colour scales with each of the independent factors of the audio stimuli, producing separate regression coefficients. Table 4.4 shows the number of significant regression coefficients found, with the positive and negative coefficients in the same cell (positive on left, negative on right). The rows in the table show the number of coefficients found below each significance level. Since two or more scales may be correlated for any subject, the same subject may appear in more than on column.

	p-level	Luminance	Spectral	Brightness	Attractiveness
Top Note	*0.05*	4/0	0/5	1/0	0/0
	0.01	0/0	0/0	1/0	0/0
	0.001	1/0	0/0	0/0	0/3
	Total	**5/0**	**0/5**	**2/0**	**0/3**
Bottom Note	*0.05*	3/1	1/1	1/1	0/0
	0.01	1/0	1/3	2/0	0/0
	0.001	2/0	0/0	0/0	0/0
	Total	**6/1**	**2/4**	**3/1**	**0/0**
Total Notes	*0.05*	0/2	0/0	1/0	1/1
	0.01	0/0	0/0	0/0	0/0
	0.001	0/0	0/0	0/0	0/0
	Total	**0/2**	**0/0**	**1/0**	**1/1**
Bottom Over Top	*0.05*	2/0	0/0	0/0	0/0
	0.01	1/0	0/0	3/0	0/0
	0.001	4/0	0/1	0/0	0/0
	Total	**7/0**	**0/1**	**3/0**	**0/0**
Top Over Bottom	*0.05*	0/0	0/0	0/0	0/0
	0.01	0/0	0/0	0/0	0/0
	0.001	1/0	0/0	0/0	0/0
	Total	**1/0**	**0/0**	**0/0**	**0/0**
Grand Totals		**19/3**	**2/10**	**9/1**	**1/4**

Table 4.4. Number of subjects with significant regression coefficients in Experiment Six for each of the main audio factors used to match with circles. Each cell has both the positive and negative regression coefficients, listed as *positive/negative*.

4.5.4 Discussion

The results in this experiment are more in line with expectations from the monochromatic results than were the results of Experiment Five. The same overall pattern of matching first-inversion chords with brighter colours is evident here as in Experiment Three. The reason that results of this experiment more effectively extend the monochromatic results than Experiment Five did lies in the nature of the stimuli. There are qualities inherent in a static chord which are more salient for most subjects in the creation of a SoC for visual comparison, when that visual stimulus varies in luminance and hue. That is, qualities of static chords are more cognitively proximal to the qualities of colour than those of a temporal interval. The work of Walker (1978) suggests that there are specific types of symbolic visual stimuli that are preferred by subjects as representative of musical constructs. Further experimentation might reveal if the cross-modal matching is more regular when notes are matched with those visual symbols that Walker designated as representing musical constructs. Some qualities of the first inversion chord, such as the transposition of the fundamental frequency (and its accompanying harmonics) to one octave higher may alter the overall perception of chordal height. Experiments One and Four demonstrate that pitch height can be regularly matched to brightness, both when colours are monochromatic or spectral.

Significant regression coefficients were observed for the number of notes in a chord. The pattern of cross-modal matching in the sample may be described as luminance differences matched most effectively with pitch height, while attractiveness was matched to a lesser degree with chord complexity (three vs. four note). The individual differences demonstrate that different subjects are sensitive to different qualities of the stimuli, and also use varying SoCs in varying alignments to perform the cross-modal evaluation and matching.

4.6 Conclusions

The experiments in this chapter demonstrate that the simple 2AFC matching task can reveal sensory correspondences when the visual stimulus is increased in complexity by the introduction of hue. Correspondences on the most basic level can be observed between luminance of presented hue and pitch height, extending the results from the previous chapter. However, when the complexity of the musical stimuli increases, so do the matching possibilities for the different auditory attributes with the multiple attributes of the visual stimuli. This revealed itself in the variety of different individual correspondences (correlations) observed for target note pitch height, interval consonance, intervallic leap direction, and size in Experiment Five. The large number of potential factors in the auditory stimuli that could be used to create a SoC for matching with the visual stimuli produced results that did not show as many same regular correspondences between modalities as the monochromatic experiments. Finally, regular correspondences were again observed in the chordal experiment, in addition to the role that attractiveness (affective response) began to play in explaining higher-level cross-modal matchings.

As the stimuli in both modalities increase in complexity, so does the composition of the SoCs that individuals may use when making cross-modal matches. Cross-modal matching tasks, when sufficiently constrained, may allow the ability to indirectly observe the salience that a particular individual places upon the different qualities of a sensory experience.

5 Cross-Modal Matching of Intervals and Timbres with Visual Stimuli

5.1 Overview

This chapter further extends the results of cross-modal matching by examining brightness and hue matching with other simple musical constructs; specifically, dyads and tones of varying timbres. The results of monochromatic brightness matching with dyads reveal that listeners remain sensitive to pitch height and use this as a SoC to align with visual brightness. Additionally, listeners begin to utilise learned musical qualities of the dyads, specifically consonance of the formed interval, as another SoC with which to form agreements. This provides evidence that such a cross-modal matching task can be used as an indirect method of observing listeners representations of higher-order musical characteristics; "eavesdropping" on such representations. The matching with timbres was performed in an attempt to determine other low-level musical characteristics that might be present in cross-modal correspondences. Synaesthetes report sensitivity to the timbres of instruments, and thus a low-level relationship between timbre and colour may exist. Cross-modal experiments do reveal an effect for timbres with missing low-order harmonics being matched with brighter colours, whether the colours are monochromatic or spectral. When the auditory stimuli vary in timbre, the tendency to match the higher-order musical characteristic of consonance observed in the first experiment is lost.

5.2 Introduction

In the previous chapter it was demonstrated that cross-modal matching was still possible with colours of a wider spectral range than those used in past monochromatic cross-modal experiments: those that varied in hue and saturation as well as brightness. When subjects were required to match these colours with single musical tones, they performed in a consistent manner matching those that had higher luminance with higher tones. Associations between higher-order musical qualities inherent in two-note temporal intervals, which were observable when monochromatic visual stimuli were used, became obscured when the visual stimuli were increased in complexity through varying the hue of the visual stimuli.

In a sense, an observable matching between sensory modalities broke down. The subjects may be forming regular matching on criteria that cannot be identified and quantified for analysis, such as semantic or affective factors. At this point, it may be informative to continue to examine more of the basic musical constructs that can affect brightness matching. Specifically, this chapter examines matching formed between dyads and monochromatic circles, and intervals composed of tones with varying timbres and monochromatic and varied hue circles.

The experiments in Chapters Three and Four examined the individual subject matching performances and demonstrated that subjects are internally consistent in matching cross-modal stimuli. These consistent matching performances may be averaged out and hence not apparent in the sample main effects, but are present nonetheless. In light of this finding, the experiments in Chapter Five will primarily examine the sample effects to see which types of SoCs most subjects use in common when matching auditory and visual stimuli. The individual variation in matching performance are still hypothesised to be present, but continued examination of these variations serves only to reiterate this finding.

5.3 Experiment Seven: Cross-modal Matching of Monochromatic Colours and Dyad Intervals

5.3.1 Aims and Introduction

Chapters Three and Four both began with experiments that examined the associations made between single tones and circles varying in their colouring. The next two experiments in both chapters examined higher order musical information being matched to the same visual stimuli. The second experiment examining the effect of a preceding note on the ratings of the subsequent tone, and the third experiment examining the effect of a chord on cross-modal matching. Both of these experiments added to the complexity of the auditory stimuli and its context. At this point, it is useful to simplify the musical stimuli again, in an attempt to clarify the role that tonality and consonance may have played in these more complex matching tasks. A better understanding of the cross-modal perception of a simple interval might lead to a better interpretation of the chordal results.

In addition, the results from Experiments Two and Four do provide us with some evidence of subjects' sensitivity to intervallic qualities. The matching was performed upon temporal intervals; however, and may be very altered when the component notes in the interval are sounded simultaneously. The supposition that a temporal interval is evaluated in the same manner as a simultaneous diatonic interval assumes a strong short-term memory component in the evaluation. That is, that the contribution of the first note is remembered and weighted as strongly in the perception of the second note as if the two were sounded together, or that any loss of first-note salience is unimportant in the perception of succeeding tones. Such an assumption must be incorrect; a note's effect on the perception of succeeding notes must decay over time. If not, an ever-increasing "echoing" of notes would rapidly create a sense of cacophony.

It is assumed that the effect of preceding notes does decay, and thus the information gained from examining a simultaneous diatonic interval may not be predicted adequately from the temporal interval conditions. Helmholtz differentiates between temporal and simultaneous intervals as being melodic and harmonic, respectively, by stating "...in melodic relationship the equality of the upper partial tones can only be perceived by *remembering* the preceding compound tone, in harmonic relationship it is determined by *immediate sensation*, by the presence or absence of beats." [Author's italics] (p. 368). Thus, if there is some different quality perceived in the *roughness* of the interval (the presence or absence of beats), then differences in cross-modal matching may also be generated.

Each stimulus in Experiments Two and Four had a fixed bass note, top note, and interval. Any of these three factors could contribute to the pattern of results. In the present experiment they were dissociated to establish which factor is most important in the cross-modal matching. These factors were the contribution of the bass note in a diatonic interval, the contribution of the top note, and the contribution of the interval type independent of constituent notes. To this end, each presented interval had a common tone, C_4, which functioned as either the bass or top note. With this restriction, all possible standard intervals available using the Western chromatic scale were presented.

5.3.2 Method

5.3.2.1 Subjects

The subjects consisted of 10 undergraduate students (6 male, 4 female) drawn from the psychology courses at the University of Stirling. Two of the subjects were musically trained, defined as having more than one year of musical training or practice on a musical instrument. None of the subjects participated in Experiments One through Six. All subjects were given credit for participation, partially fulfilling a degree requirement.

5.3.2.2 Equipment

Visual stimuli were presented on a NeXTstation TurboColor computer, equipped with a 19" VDU. Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a Yamaha DX-7 Synthesizer using its standard *Harp-Flute* patch. The synthesizer's output was synchronised with the NeXT computer by means of a standard Midi connection. The subject was seated at a comfortable viewing distance from the VDU (50 centimetres), subtending a 29° visual angle. The subject provided input to the NeXT by means of the two-button mouse. This equipment setup was identical to the setup in Experiment One.

5.3.2.3 Stimuli

5.3.2.3.1 Intervals

Twenty-four separate dyad intervals, each with C_4 as a common tone were used. These were subdivided into two groups of 12 intervals, forming the basic intervals of the Western tonal scale: m2, M2, m3, M3, P4, aug4, P5, m6, M6, m7, M7, P8. Each interval was presented a total of five times. Figure 5.1 shows the interval stimulus set in musical manuscript representation.

Figure 5.1. Intervals used as stimuli in Experiment Seven. Although the notes are written separate, in the actual experiment the two notes in a bar are sounded simultaneously.

5.3.2.3.2 Visual

Two circles varying in grey levels as in Experiment One (see Section 3.2.2.3.2).

5.3.2.4 Procedure

The subjects were told that they would hear a tone through the headphones while two circles would simultaneously appear on the computer monitor. They were instructed "...to select the circle that you feel is most like the tone. If you feel that the one on the left is more like the tone, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of tone and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The subjects were played each of the 24 intervals five times each for a total of 120 stimulus intervals. These tones were presented in a random order. The program randomly selected a grey level for one of the two circles that ranged in grey value for 0.1 to 0.5. The selection of the one circle determined the grey level of the second circle, as the second circle was offset by a constant difference of 0.4 for discriminability, as described above. The left or right positioning of the two circles was also randomised on each trial. The subject was thus provided with forty-eight separate trials of tone and circle pairs. After the last trial, the screen went blank, and a "Thank You" message appeared.

5.3.3 Results

Scores for each interval were computed as the number of times that the brighter of the two circles was chosen when presented with an interval stimulus. Thus, a score of "4" indicates that the brighter grey circle was chosen four times out of five for that interval. A two-way repeated measures ANOVA was performed for the direction of formed interval (*Up*, with C_4 as the bass note of the interval; *Down*, with C_4 as the top note) and the twelve basic diatonic interval types. Significant main effects are observed for the direction of interval ($F(1, 9) = 14.248$; $p < .01$), diatonic interval type ($F(11, 99) = 5.285$; $p < .001$), as well as a significant interaction between these two independent variables ($F(11, 99) = 3.564$; $p < .001$). Intervals with C_4 as the bass note had a mean score of 3.392; those with C_4 as top note were matched with darker circles, and had a mean score of 1.417. Figure 5.2 plots the scores against the twelve different intervals. Figure 5.3 is identical in layout, but separately plots those scores for the intervals with C_4 as bass note and top note representing the interaction between interval type and direction of interval.

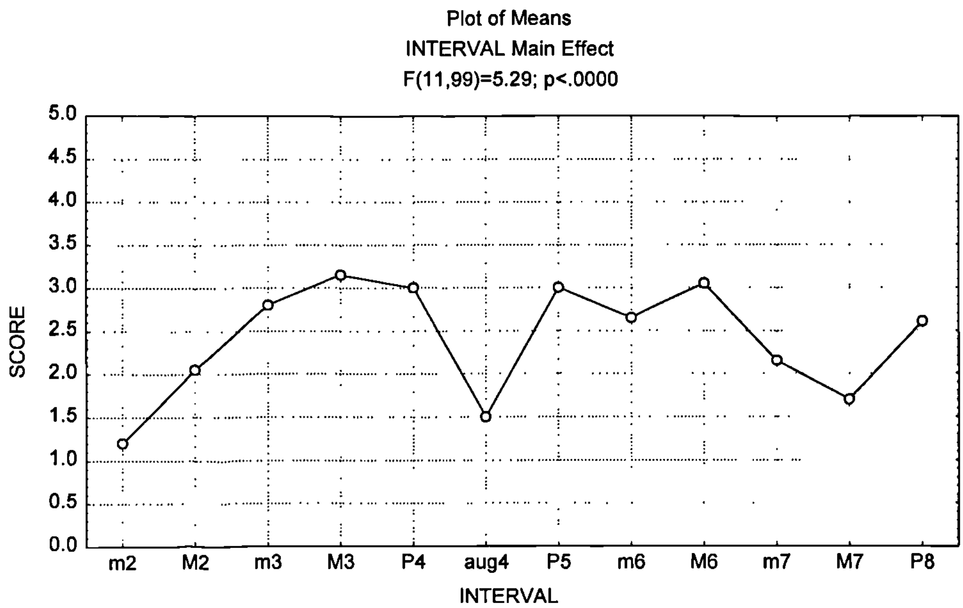


Figure 5.2. Scores for each diatonic interval collapsed across interval direction.

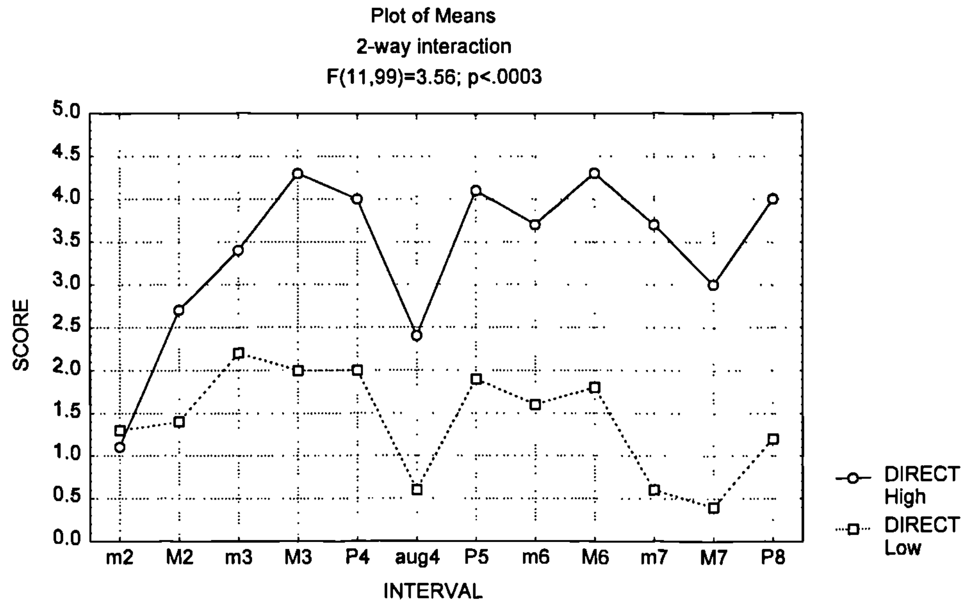


Figure 5.3. 2-way interaction plot of scores plotted against interval type.

The matching performance with intervals can also be analysed by using consonance/dissonance (referred to hereafter as *consonance*) as an independent variable, as was done in Experiments Two and Five. Helmholtz proposes the following different ratings of consonance based upon roughness of constituent harmonics (in descending order or consonance):

- Absolute consonances: Unison (not used) and Octave (P8)
- Perfect consonances: P4 & P5
- Medial consonances: M6 & M3
- Imperfect consonances: m3 & m6

The remaining intervals of m2, M2, m7, M7, and aug 4 were not named, but referred to as *dissonant intervals* throughout (p. 194). For the purposes of analysis, these intervals can be grouped into two groups of six, if we include the interval of m6 with the dissonant interval. Helmholtz suggests that the characteristics of this interval often makes it more dissonant than a minor 7th, due to its roughness and proximity to the second most consonant interval of a Perfect 5th.

When the data from the experiment are examined with this additional IV, it is possible to perform a 2 x 2 x 6 repeated measures ANOVA on the data using direction of interval, consonance of formed interval, and the six separate intervals in the stimulus set. Main effects were are again observed for the direction of the interval (F(1, 9) = 14.248; p < .01) as well as a main effect for the consonance of the formed interval (F(5, 25) = 4.843; p < .01). The main effect for the interval type is no longer present when consonance is removed as a factor. Two-way interactions are observed between direction and interval (F(5, 45) = 3.142; p < .05) and consonance and interval (F(5, 45) = 4.843; p < .01), as well as a three-way interaction between interval, direction, and consonance (F(5, 45) = 4.614; p < .01). Figure 5.4 plots the two-way interaction of interval and consonance.

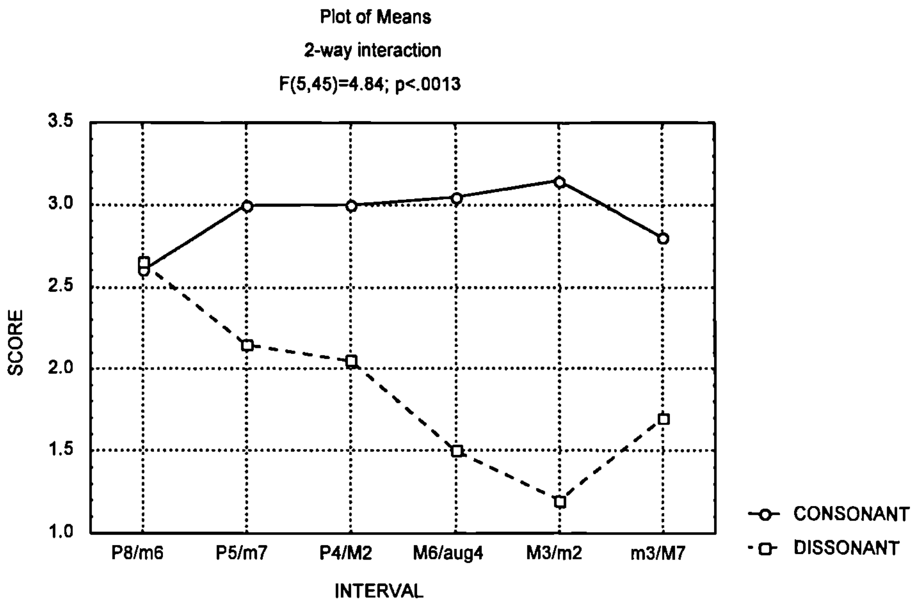


Figure 5.4. Two-way interaction of scores for interval type and the consonance of the interval.

5.3.4 Discussion

The results demonstrate that cross-modal matching performance is different for the dyad intervals (two notes sounded simultaneously) and temporal intervals (two notes sounded serially). Brighter circles are still cross-modally matched with the intervals of increasing pitch height; in this experiment, that pitch height is composite of the two constituent notes. The term *composite pitch height* is used to designate the pitch height as a function of the non-C_4 dyad note. All stimulus intervals have the note of C_4 in common, but the second note of the interval alters the overall interval pitch height. That is, the "high" intervals all have the second note higher, and consequently have many higher audible partials than the "low" intervals. The finding of increasing composite pitch height aligned with increasing brightness follows naturally on from the past experiments. What would not have necessarily been predicted from the temporal interval experiments (Two and Four) is the distinctive pattern of cross-modal matching between different interval types in a musical sense and visual brightness.

The most notable feature in both Figure 5.2 and Figure 5.3 is the much lower scores of brightness matchings with the augmented fourth, or *tritone* interval. The tritone figures prominently as an interval with many rules surrounding its use in classical music theory. Leaps of a tritone are to be avoided in melody lines, and its use in a chord or between melodic voices in a composition necessitates following specific rules to achieve its resolution. This may be due to the physical nature of the interval, as a great many of the two notes' audible harmonics are sufficiently close to create beats, and none overlap to reinforce each other. This quality of harmonics creating beats contribute to what some researchers refer to as the interval or chord's *roughness*. Although not a strict definition for a musical theorist, this quality has been proposed to be the psychophysical basis of the musical concept of *dissonance*. Helmholtz concisely describes the qualities as being one and the same:

> *When two musical tones are sounded at the same time, their united sound is generally disturbed by the beats of the upper partials, so that a greater or less part of the whole mass of sound is broken up into pulses of tone, and the joint effect is rough. This relation is called* Dissonance.

> *But there are certain determinate ratios between pitch numbers, for which this rule suffers an exception, and either no beats at all are formed, or at least only such as have so little intensity that they produce no unpleasant disturbance of the united sound. These exceptional cases are called* Consonances. (pg. 194)

In the case of the tritone, the dissonance is so extreme and distinctive that some early theorists referred to it as *diabolus in musica*, or "the devil in music." (Cooke, 1959, p. 43). Whatever quality the tritone possesses, it is apparent that subjects matched it with darker circles than other intervals. If this is so, subjects were then demonstrating sensitivity to roughness or dissonance, and using these qualities in a SoC for cross-modal matching to monochromatic brightness, in addition to use of overall interval pitch height.

Another possible explanation for the change in matching performance may lie in the so-called *tritone-paradox*, described by Deutsch (1986). When presented with a temporal tritone interval using complex tones (composed of many sine-waves with a Gaussian amplitude envelope), subjects showed confusion as to whether the second note was higher or lower than the first. In essence, the perception of the interval could change so that either of the two notes were perceived alternately as the top or bass note of the interval. It is possible that pitch height confusion proceeds in part from a unique quality of the tritone, and that the cross-modal matching performance reflects this confusion.

These results of the matching also strongly indicate that subjects were sensitive to consonance of the interval in addition to the interval's overall pitch height in their cross-modal matching. The interaction of interval type in all three significant interactions suggests that subjects may be sensitive to other characteristics of the intervals. The decreasing scores in Figure 5.4 for the dissonant intervals suggests that perceived "brightness" of the intervals decreases as the interval becomes increasingly dissonant, as these intervals are approximately ordered along the abscissa as a function of increasing dissonance. No such pattern is evident in the ratings of the six consonant intervals, which are also placed along the abscissa as a function of increasing consonance. A planned-comparison between the three most (M7, m2, aug4) and second most dissonant intervals (M2, m7, m6) does in fact reveal a significant difference in cross-modal brightness matching scores ($F(1, 9) = 21.849$; $p < .01$), where no such difference is present in the three most and second most consonant intervals. One explanation for this finding is a potential threshold in dissonance that must be exceeded before consonance becomes important in matching. While the interval is consonant, or *agreeable* in affective terms, the matching is performed by aligning the pitch height SoC for the interval with the brightness SoC for the circles. When the interval increases in dissonance, this factor becomes more salient in the perception of the interval, and the auditory SoC incorporates both the interval pitch height and dissonance of the interval.

The strong association between visual brightness and interval consonance further brings the proposed equivalence of pitch height and colour brightness as some central representation into question. More likely the use of visual brightness as a SoC for matching tasks and its use in experiments provides significant relationships with other modalities because it can be employed in a consistent fashion. These results also suggest that cross-modal matching techniques can be of use in an instructional context. If students can spontaneously match musical material with visual material in a regular fashion, use of the visual information may draw attention to the attribute of the music that an instructor may be attempting to explain. Experiment Eight will examine this possibility in more detail.

5.4 Experiment Eight: Cross-modal Matching of Monochromatic Colours and Dyad Intervals of Varying Timbres

5.4.1 Aims and Introduction

Experiment Eight continues to examine cross-modal matching with chords by simplifying the components of the complex tone itself. That is, it examined the subjects' responses to different intervals that also differ in the timbre of their constituent complex tones.

Synaesthetic research has attempted to draw conclusions about cross-modal associations between vowel sounds and qualities of evoked colours or photisms (including brightness and/or hue). The perceived pitch and height of a spoken vowel is dictated by the fundamental and second formant of the vowel (Marks, 1975). Given this, modification of the complexity of a tone should also alter the cross-modal matching of circle brightness with dyad interval. In the same manner as the earlier experiments, subjects are not expected to actually respond to these musical factors in precisely the same manner, although consistency in matching should be observable as long as the different qualities of the dyads are discriminable. This experiment examined the role that the timbre plays in isolation along with the potential ways that it interacts with the overall pitch height and interval type of the musical stimulus.

This experiment also began to examine the role of screen background on which different visual stimuli are presented. Little cross-modal research has been done on the potential role that visual background plays upon the matching of visual information with audio. Synaesthetes describing their perception often omit a description of the background on which their visual perceptions might be placed. One subject, AL (see Chapters Seven and Eight for details), found it impossible to describe any sense of background when pressed to describe it. Since it is impossible to present visual stimuli without some background, the effect that different backgrounds may play in cross-modal correspondences should be considered at least in general terms.

In order to look at the potential effects that timbre and background might play in cross-modal judgments, it was assumed that the overall pitch height of the intervals would play at least a significant role in the matchings, as suggested in Experiments One, Four, and Seven. A combination of all of the various independent factors would create an experiment too lengthy to be performed. In an effort to avoid this, the bottom note of the stimulus interval was allowed to be selected purely at random between the range of D_2 to Ef_5. Using this design, it was still possible to examine the effect that the interval pitch height plays upon brightness matching. This design also allowed for the examination of the separate roles that bottom and top note play in the brightness matching.

5.4.2 Method

5.4.2.1 Subjects

The subjects consisted of 10 undergraduate students (4 male, 6 female) drawn from the psychology courses at the University of Stirling. Two of the subjects were musically trained, defined as having more than one year of musical training or practice on a musical instrument. All subjects were given credit for participation, partially fulfilling a degree requirement. None of the subjects had participated in any of the preceding experiments.

5.4.2.2 Equipment

Visual stimuli were presented on a NeXTstation TurboColor computer, equipped with a 19" VDU. Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a NeXTstation's digital signal processor (DSP56001). The subject was seated at a comfortable viewing distance from the VDU (50 centimetres), subtending a 29° visual angle. The subject provided input to the NeXT by means of the two-button mouse.

5.4.2.3 Stimuli

5.4.2.3.1 Auditory

The twelve intervals of the Western tonal scale: m2, M2, m3, M3, P4, aug4, P5, m6, M6, m7, M7, P8. Additionally, a note in isolation is presented (unison). The base note of the interval randomly varied from A_2-Ef_5, and the top notes varied from A_2 (in the unison) to Ef_5. The number of trials at any base note was not controlled; instead, each interval was presented 5 times for each timbre.

5.4.2.3.2 Visual

Two grey circles were used for the visual stimuli, as in Experiment One. The two circles varied in grey level from 0.1 to 0.9 (postscript), each two separated by an interval of +0.4. The circles were placed on either a completely white background or a completely black background (see Figure 5.5).

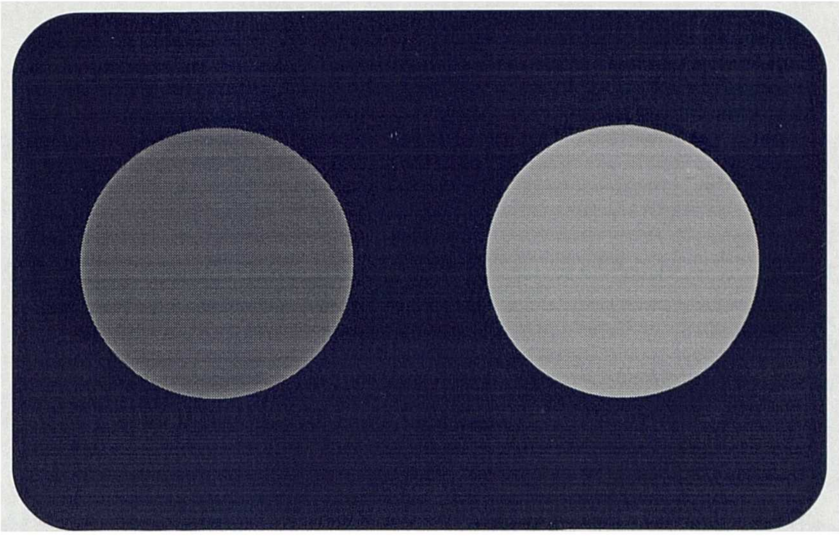


Figure 5.5. Sample monochromatic stimuli presented on a black background for Experiment Eight.

5.4.2.3.3 Timbres

The timbres of the waveforms were created with the NeXTstation's DSP. The different timbres were composed of simple sine waves, with no attack or decay envelopes. The different timbres were created by adding different simple sine waves for the stimulus note's harmonics. These constituent waves may or may not include the stimulus note's fundamental frequency. The harmonics used in the creation of these timbres are as follows:

1. $H_0 + H_1$;
2. $H_0 + H_1 + H_2 + H_3$;
3. $H_2 + H_3$.

The amplitudes of the constituent sine waves decreased in a manner that roughly modeled the harmonic sequence of a musical tone. Amplitudes of the harmonics are expressed as percentages of the fundamental amplitude. The amplitudes of the harmonics sine waves for H_0, H_1, H_2, & H_3 were 100%, 75%, 50%, and 25%, respectively. Figure 5.6 plots the power spectrum for each of the three experimental timbres. The percentage of the fundamental frequency is shown on the ordinate with the four different harmonics located in order along the abscissa.

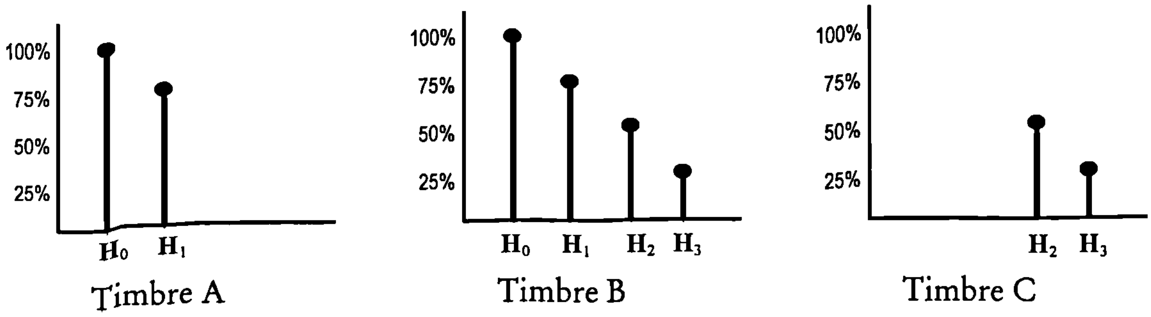


Figure 5.6. Power spectra for the three experimental timbres used in experiments Eight and Nine.

5.4.2.4 Procedure

The subjects were told that they would hear a tone through the headphones while two circles would simultaneously appear on the computer monitor. They were instructed "...to select the circle that you feel is most like the tone. If you feel that the one on the left is more like the tone, press the left mouse button. If you feel that it is more like the one on the right, press the right mouse button. After making your selection, press that mouse button again, and the computer will give you the next set of tones and circles. Please continue until the computer ceases giving you the pairs." At this point, any questions were answered, and the experiment commenced. The subjects were played each of the 13 separate intervals using each of the 3 timbres. This created a stimulus set of 39 total intervals. Each of these intervals was presented 6 times, with the bass note of the interval randomly selected from A_2 to Ef_5, creating a total stimulus set of 234 individual intervals. Since the largest stimulus interval was an octave (P8), the top note of the stimulus set was an Ef_6. These intervals were presented in a random order. The program randomly selected a grey level for one of the two circles that ranged in grey value for 0.1 to 0.5. The selection of the one circle determined the grey level of the second circle, as the second circle was offset by a constant difference of +0.4 for discriminability, as described above. The left or right positioning of the two circles was also randomised on each trial. In addition, the circles were presented on either a white or black background. The background was randomly selected for each trial. The subject was thus provided with 234 separate trials of tone and circle pairs. After the last trial, the screen went blank, and a "Thank You" message appeared.

5.4.3 Results

For the initial analysis, scores were left in binary format, specifying whether the darker or brighter (0 or 1, respectively) of the two presented circles were chosen for a particular trial. Highly significant main effects are observed for the interval pitch height, both when analysed by the interval's bass note height ($F(29, 2310) = 11.66$; $p < 0.001$) or the interval top note height ($F(41, 2298) = 8.10$; $p < 0.001$). Figure 5.7 plots the mean scores for brighter circle matched with intervals when analysed by bass note height (left plot) and top note height (right plot).

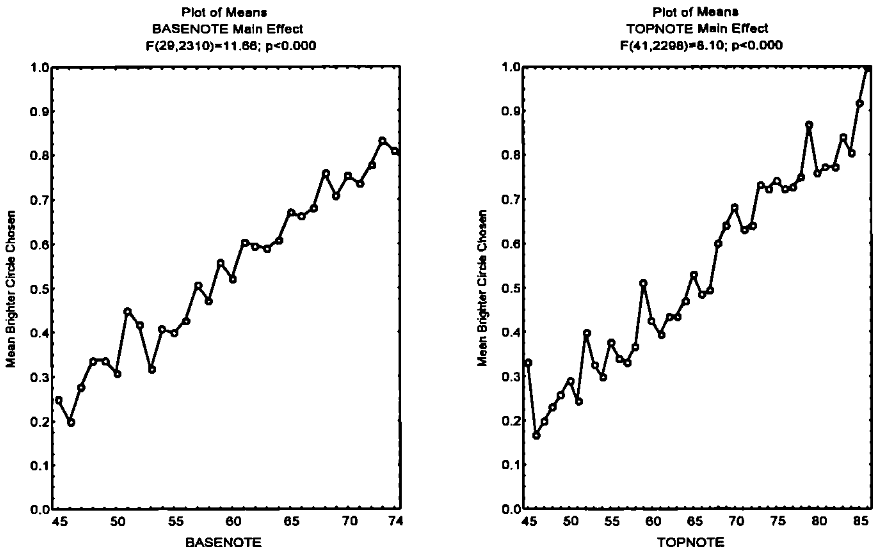


Figure 5.7. Means scores for brighter circle chosen for musical intervals when categorised by intervallic bass note (left) and top note (right). The numbers along the abscissa indicate the pitch height of the bass note or the top note of the stimulus interval.

A 2 x 3 x 13 repeated-measures ANOVA examining the factors of background (black and white), timbre (A, B, & C), and interval type (Unison and 12 basic diatonic intervals) was run on the brightness scores for each stimulus. A significant main effect is observed for background colour ($F(1, 2262) = 19.92$; $p < .001$), with subjects selecting the brighter circle more often on those trials presented on a black background (Mean = 0.577; SE = 0.014) than when on a white background (Mean = 0.489; SE = 0.015). Significant main effects are also observed for timbre ($F(2, 2262) = 94.47$; $p < 0.001$) and interval type ($F(12, 2262) = 1.77$; $p < 0.05$). No interactions are observed between the independent variables. Figure 5.8 plots the mean scores for brighter circle chosen against timbre type. Figure 5.9 plots the mean score for brighter circle chosen against interval type. Figure 5.10 plots the means for brighter circle chosen for the different intervals with separate lines for each timbre. Separate graphs are provided for each background colour. No systematic differences are observed between perfect, consonant, and dissonant intervals.

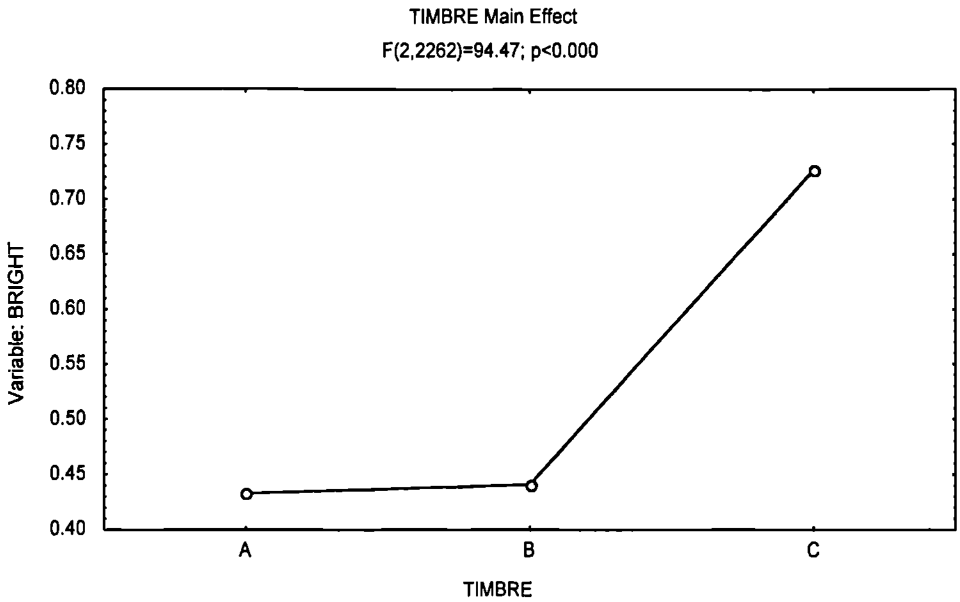


Figure 5.8. Mean scores for brighter circle, with means collapsed across interval type and background colour.

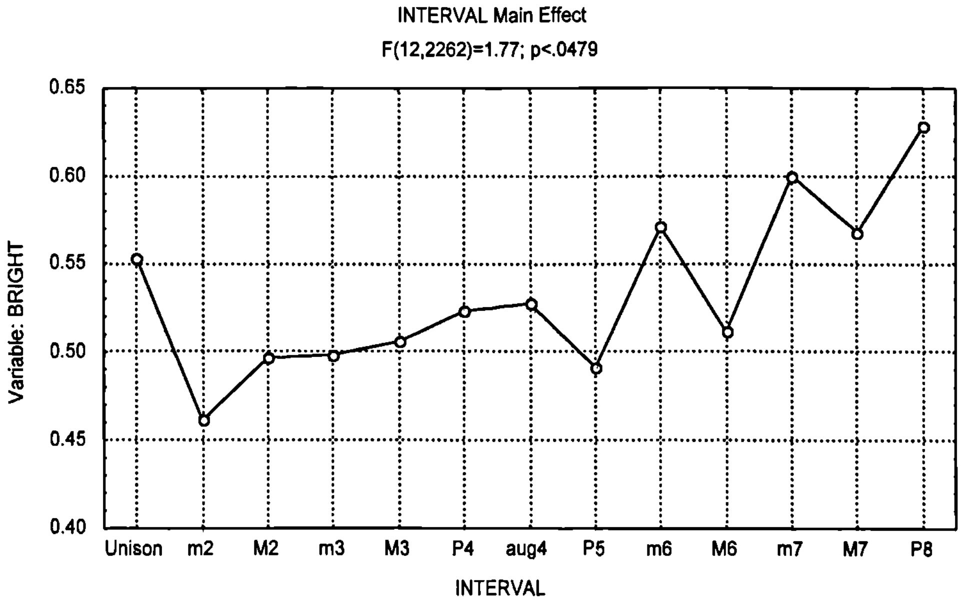


Figure 5.9. Mean scores for brighter circle chosen for all interval types, collapsed across timbres and background colour.

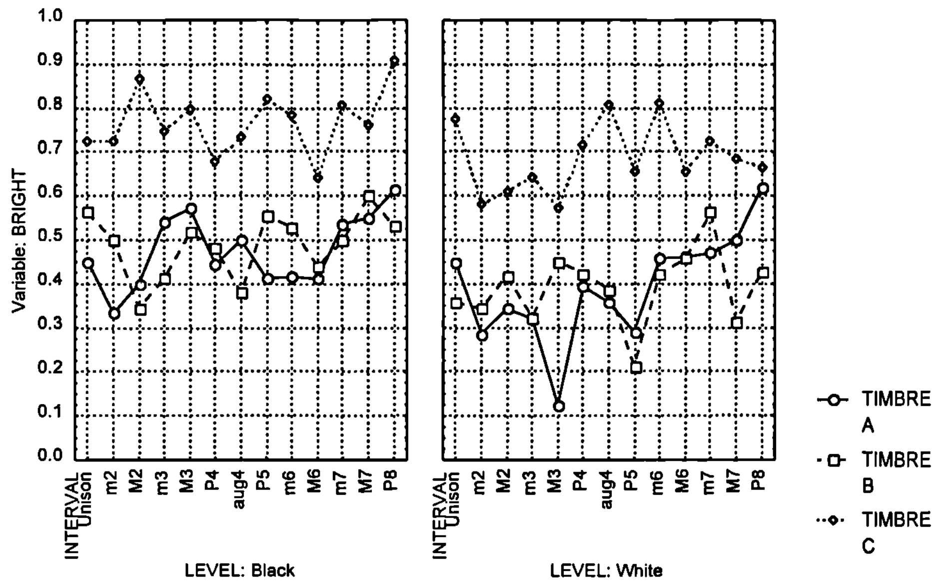


Figure 5.10. Mean scores for brighter circle chosen for intervals and timbres, plotted separately for different background colours.

5.4.4 Discussion

The results of Experiment Eight further demonstrate the strong role that overall pitch height of a tone or interval plays in cross-modal brightness matching. One of the most interesting additional characteristics of Experiment Eight results was the strong role that timbre plays in the cross-modal matching using different intervals. When examining the main effect of timbre in visual matching, no obvious differences are noted between Timbres A & B. That is, the addition of the two higher level harmonics did not alter the matching patterns; at least, not strongly enough to override other factors in the SoCs that are used for brightness comparisons. However, the elimination of the fundamental and first harmonic, in Timbre C, causes an abrupt increase in the number of brighter circles chosen to match the interval. The perception of a complex tone as a pitch relies on a variety of auditory aspects. One feature of pitch perception is the diversity of harmonic envelopes that are perceived as the same pitch. It has been well established that the absence of a tone's low-order harmonics does not alter perception of the pitch itself. This effect is sometimes referred to as the *residue phenomenon* (Dowling & Harwood, 1986; pp. 27-28). In pilot trials for Experiment Eight, subjects did report that stimuli with a fundamental frequency of 220 hertz and produced with Timbres A, B, & C were each perceived as an A_3, although the different timbres were different in character.

The results of the timbre matching experiment suggest that the absence of the lower frequencies in the timbre plays a more important role in the SoC for cross-modal matching than the addition of the higher frequencies. The analysis of cross-modal matching using interval bass note pitch height further supports this: intervals with lower bass notes necessarily have lower frequencies, and are matched with darker circles. Of course, analysis by top-note also reveals a significant positive trend in brightness ratings, but these top notes are well-correlated with the base-notes by virtue of the experimental design. These results follow on logically from the results of the chordal experiments, which found chords in first inversion matched with brighter circles.

One remaining question is what has happened to the effect of consonance, evident in the results of Experiment Seven but absent here. The answer lies in the salience of factors used to compose a SoC for a particular experiment. In Experiment Seven, the musical factors of pitch height and consonance are the most apparent to the listener, as most other factors have been controlled. When creating a SoC for comparison with the circles, those factors are employed by the listener, as no others are available. In Experiment Eight, those same factors are available for the subject to use; however, the subject now has the factor of varying timbre available for use in a SoC. The results suggest that timbre differences are closer in cognitive proximity to grey level differences than consonance. Timbre overrides consonance as a salient factor in the experimental task. Again, the choice of stimuli has a direct relationship on the different SoC used for the matching task.

5.5 Experiment Nine: Cross-modal Matching of Colours and Dyad Intervals of Varying Timbres

5.5.1 Aims and Introduction

In light of the positive results of Experiments Seven and Eight, it remains to be seen if the results change when adding varying hue to the monochromatic visual stimuli. The monochromatic single tone and chord experiments did reveal similar patterns of results when the circles increased in complexity by varying their hues. Similar patterns of results were hypothesised for this experiment, as the dyads and varying timbres are more similar to chords than temporal intervals. This experiment specifically examined whether Timbre C from Experiment Eight, with the missing fundamental and first harmonic, yield cross-modal matching with more luminous circles as previously observed. Matching of more luminous circles with higher overall interval pitch height was also predicted. The overall pitch height was hypothesised as still being the most salient quality of the auditory stimuli, with timbre factors adding to the creation of SoCs for cross-modal matching.

5.5.2 Method

5.5.2.1 Subjects

The subjects consisted of 10 undergraduate students (4 male, 6 female) drawn from the psychology courses at the University of Stirling. Two of the subjects were musically trained, defined as having more than one year of musical training or practice on a musical instrument. All subjects were given credit for participation, partially fulfilling a degree requirement. All of the subjects participated in Experiment Eight.

5.5.2.2 Equipment

The equipment was identical to the equipment from Experiments One through Eight.

5.5.2.3 Stimuli

5.5.2.3.1 Intervals

The twelve intervals of the Western tonal scale: m2, M2, m3, M3, P4, aug4, P5, m6, M6, m7, M7, P8. Additionally, a note in isolation is presented (Unison). The base note of the interval randomly varied from A_2–D_5, and the top notes varied from A_2 (in the unison) to D_6. The number of trials at any base note was not controlled; instead, each interval was presented 5 times for each timbre.

5.5.2.3.2 Visual

Two coloured circles varying hue, brightness, and saturation. The exact levels of each of these colours on the NeXT display HSB model are provided in Table 4.1. The colours were divided into three sets; one set at full brightness and saturation, one at 50% brightness and full saturation, and one at 50% saturation and full brightness. Each set was composed of six circles, with the colours in each set taken from the primary and secondary hues: red, orange, yellow, blue, green, violet. Thus, each of these six hues was presented at three brightness/saturation levels, comprising a total set of 18 distinct colours. Because of the great number of pairings that would be possible with 18 colours (152), and to control the presentations for controlled judgments made by varying hue, each colour was presented with another colour from its same saturation/brightness group.

5.5.2.3.3 Timbres

- The timbres of the waveforms were identical to those used in Experiment Eight, and are described in Section 5.5.2.3.3.

5.5.2.4 Procedure

Subjects heard the intervals and saw coloured circles appear with the onset of the interval. The subject was asked to choose the circle that was *more like* the interval. The procedure is identical to that for Experiment Eight.

5.5.3 Results

For the initial analysis, scores were left in binary format, specifying whether the darker or brighter in terms of cd/m^2 (0 or 1, respectively) of the two presented circles was chosen for a particular trial. One-way ANOVA reveals no significant main effect for interval pitch height and visual brightness when computed for either interval bass note height or top note height. A one-way ANOVA performed on the scores of more luminous circle chosen and the three timbres does reveal a similar main effect as found in Experiment Eight ($F(2, 2247) = 9.42$; $p < .0001$). A plot of the means for the timbres in shown as Figure 5.11.

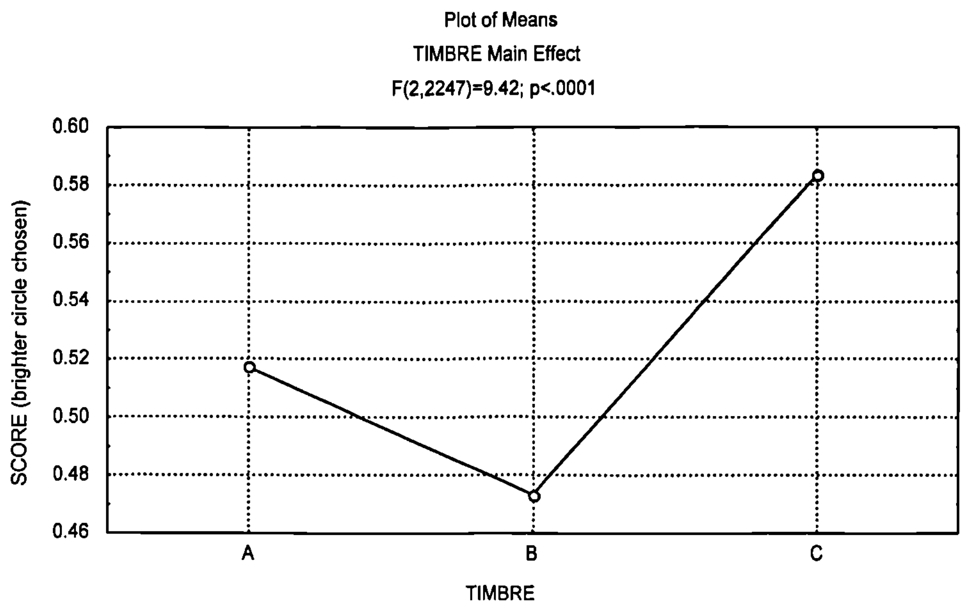


Figure 5.11. Mean scores for more luminous circle chosen for the different timbres, with means collapsed across interval type and background colour.

A 3 x 13 repeated-measures ANOVA examining the factors of timbre (A, B, & C) and interval type (Unison and 12 basic intervals) was run on the scores of more luminous circle chosen for each auditory stimulus. The same main effect is observed for timbre, but no main effect for interval type or interaction between factors was observed. A one-way ANOVA performed on the scores and the consonance of intervals does reveal a significant effect ($F(1, 2248) = 4.983$, $p < .05$), with more dissonant intervals being matched with more luminous circles.

5.5.4 Discussion

The regular associations observed in Experiment Eight begin to break down with the introduction of hue to the visual stimuli, in the same manner that sample effects were reduced in the results of Experiment Five. The circles varying in hue presents the listener with many potential qualities to use as a SoC, generating idiosyncratic matching performance for many of these qualities. This pattern is similar to the results of Experiment Five.

The relationship between pitch height and brighter or more luminous circle chosen, so strong in previous experiments, is no longer present in this experiment. However, matching between Timbre C and more luminous circles is still present, as predicted from the results of Experiment Eight

This provides more evidence that visual and auditory brightness need not be represented in a single suprasense. The hypothesis that pitch height is equated with brightness is seen again to be overly simplistic. When presented with varying timbres, subjects favour aligning the SoC of colour brightness with timbres, rather than using the qualities of overall pitch height. Thus, there is some quality of tones with absent fundamentals that align more naturally with the brighter pole of the colour SoC.

Correlations run on each subject's response score revealed that no subject demonstrated a significant alignment of timbre and luminance, although the sample did. The main effect of matching interval consonance with colour luminance suggests that this factor was used in SoCs for subjects, and accounts for much of the residual variance in the data. Contrary to the results of Experiment Seven, the direction of association was now that the more dissonant intervals were matched with brighter circles than the consonant intervals. Again, the nature of the stimuli defines the SoCs and alignments used by subjects in a cross-modal matching task.

5.6 Conclusions

Cross-modal matching using dyads and monochromatic circles generated regular matching patterns in Experiment Seven. Subjects demonstrated the ability to use musical qualities, such as consonance, as a SoC for cross-modal matching. When the auditory stimuli were further varied, by using different timbres, subjects did not use the consonance information in an observable manner. Instead, the varying timbre became an important SoC. It can then be said that timbre differences are more closely related or salient—have a closer cognitive distance—to dyads of varying timbre than consonance differences.

Subjects also demonstrated that the absence of lower frequencies in dyads was an important SoC for cross-modal matching when presented with notes of varying timbres. Dyads with less energy in the lower auditory frequencies were matched with brighter circles. This correspondence between absence of lower frequencies and brighter visual matching is supported by the chordal results of first-inversion chords matched with brighter circles described in earlier chapters. When the visual stimuli were allowed to vary in hue, the matching performance again changed. Pitch height then ceases to be an important factor in cross-modal matching, although the varying timbres remained important in SoC creation.

The results of the three above experiments demonstrate that cross-modal matching between visual brightness and musical tones is not always primarily governed by tone height. Listeners in these experiments might form cross-modal matches by only aligning composite pitch height or another musical quality (such as consonance) of dyads with varying brightness. If a low-level direct relationship existed between visual and auditory brightness, the willingness to use brightness as a SoC to compare with musical qualities would not be anticipated. Instead, it has been demonstrated that observing brightness matching with notes may indicate which integral dimensions of the auditory stimuli are most salient given the matching task's parameters. Brightness is matched with tone pitch height primarily when tones have constant timbres. However, in the presence of varying timbres, the salience of the dimensions of timbre quality and pitch height may both play a part in the forming of the auditory SoC that is aligned with varying brightness.

The results also provide additional evidence that care must be taken when hypothesising sensory characteristics that are present within multiple modalities. While the relationship between the percepts in testing may appear to be completely regular, other percept alignments not yet investigated may be equally regular, and given preference in matching tasks. In the case of Experiment Seven, varying musical stimuli by combining tones into dyads may have increased the number of integral dimensions, but perceived pitch class remained regularly aligned with brightness. However, subjects did use information from those additional dimensions in their matching performance. Thus, the association was mediated by a more complex auditory SoC. When the stimuli increase in variability through the controlled alteration of timbres, another alignment became apparent: listeners were sensitive to differences in timbre and demonstrated a strong tendency to align timbres that had missing low-order harmonics with brighter colours. When the variability in colours was then increased by adding a varying hue component, the alignment between timbre and brightness remained at the expense of a pitch height–brightness alignment.

The internal representation of a musical element as basic as a single tone is complex. Tone timbre, pitch class, harmonics, and waveform transients (onset and offset) are all available for encoding as perceptual qualities of an individual note. When a second tone is added, either sounded simultaneously (dyad) or offset in time (two-note temporal interval), the complexity of representing the experience increases dramatically. The results of the experiments in Chapters Three, Four, and Five show that subjects combine many of these factors into SoCs when requested to draw cross-modal matches. The nature of the task dictates the type of SoC created, and the alignment of that SoC with the scale it is being compared to, or *source scale*. When the source scale is simple, such as the monochromatic circles that vary only in grey level, a SoC composed of more than one dimension may still be translated onto that scale, and subjects can consistently make such matches.

Something inherent in a brightness SoC aligns naturally with the auditory qualities of tone height and timbre: colour brightness and luminance are cognitively proximal to auditory qualities of the stimuli. Such alignments are not the sole comparisons that can be used with such visual scales: the SoCs of colour brightness and luminance may also be matched with higher-order musical factors of consonance and direction of intervallic leap. Thus, the regular matching between percepts from multiple modalities demonstrates an active, dynamic process of choosing scales to draw sensory matchings. Individuals actively seek to form regular correspondences regardless of integral dimensions. Some factors which may aid in the alignment of these SoCs are affective similarities and sensory similarities (*e.g.*, magnitude and duration).

6 Cross-Modal Matching of Higher-Order Musical Passages with Adjectives and Landscapes

6.1 Overview

This chapter examines subject matching of Gaelic folktunes played on a computer with adjectives pairs and pictorial landscapes. Subjects initially rated all tunes on a linear scale representing the affective scale of *Very Happy* to *Very Sad*. High levels of agreement for affective ratings were observed for a subset of the tunes. The four happiest and four saddest rated tunes were then used as a stimulus group to examine affective ratings on short passages of varying beats either beginning on the first note of the tune or concluding on the final note. High levels of agreement were again observed, and provided an indication of some musical constructs that contribute to affective ratings, such as large intervallic leaps increasing happiness ratings of fragments. Each of the Gaelic folktunes was also evaluated using a modified *semantic differential* task, using only adjectives representing poles of a semantic scale (*e.g.*, *Warm/Cold*) and forcing the listener to match one of the adjectives with the piece. Those pieces that were rated among the happiest and saddest in the first experiment were also among those pieces that had the most agreements in adjective selection. Factor analysis revealed that many of the adjective pairs redundantly described two main semantic factors along which subjects represented the music. These results were extended by using landscape pictures rather than adjective pairs for matching. The four saddest rated tunes from Experiment Ten were among the five tunes with highest picture selection agreement. Factor analysis revealed that the use of pictures increased the number of representative factors to four, demonstrating that the tunes may be represented along more dimensions than may be expressed with language.

6.2 Introduction

Chapters Three, Four, and Five examined cross-modal comparisons and matchings of basic auditory and visual perceptual qualities; specifically, colours and simple musical elements. It was concluded that agreements need not be based upon common suprasensory representations, but could be produced through alignment of different *scales of comparison* (SoCs). These scales were hypothesised as evolving from existing scales. Scales were said to have *cognitive distance* from each other, defined as how closely they are related from the course of development. Scales that are cognitively proximal for most individuals will produce patterns of matchings that are similar in nature, and observable in cross-modal experiments.

The experiments in this chapter now examine the matches made between high-level stimuli. The previous experiments all used the auditory basic elements of music as stimuli: these will use complete musical melodies. The increase in complexity is large, but using this type of stimuli can reveal if the SoC hypothesis can work for stimuli of a much greater complexity. The other scales presented for evaluating the music will also be increased in complexity: affective scales and adjective scales.

The process of comparing and matching different modal-specific evaluations has a correlate in evaluating the semantic components inherent in most sensory perceptions. Coloured-music synaesthetes report responding to a variety of different aspects of music: the colour of a particular note, a chord, the timbre of a particular instrument, cadences, or indeed the colour of an entire movement. If it is possible for synaesthetes to align complex musical qualities with colour, it may also be possible for normal subjects to match entire passages of music to multiple SoCs. The experiments in this chapter will examine such matchings.

The hypothesis that concepts occupy positions in an undefined semantic space has been discussed for decades (Osgood, Suci, & Tannenbaum, 1957). Osgood et al examined responses to different words utilising their technique of *semantic differentials*. This method requires the subject to rate a single word on several predetermined categories that comprise adjective continua, (*e.g., Happy* to *Sad*). Using this technique, the experimenters identified an approximate "position in cognitive space" for the word or concept. However, it is necessary to point out that these dimensional scales are identified by *the experimenter*, and may or may not correctly identify scales of comparison that the subject utilises in the formation of a conceptual representation.

It is also entirely possible that several of these supplied adjective dimensions correlate so well as to be redundant in describing the evaluative process. However, using many scales is desirable, as different subjects or cultural groups will have evolved different scales, and the saliency of each of these scales will vary amongst the population.

6.3 Experiment Ten: Happy/Sad Ratings of Gaelic Folk Tunes

6.3.1 Aims and Introduction

Experiment Ten sought to investigate the reliability with which a simple unaccompanied tune communicates a "sense" that projects onto an evaluative scale of happiness. That is, it attempted to determine if subjects could reliably use an adjective scale ranging from *Happy* to *Sad*. To accomplish this, subjects were asked to evaluate how happy or sad they felt the particular melody was. The stimuli were chosen from a collection of unaccompanied gaelic folk tunes (MacPharlain, 1908). The tunes are largely unknown, and thus devoid of previous associations except for those individuals well versed in the style. In fact, only one subject recognised any of the tunes, and this subject only recognised a single tune. The monophonic nature of the melodies allowed for the examination of simple constructs such as tempo, syncopation, and implied tonality without the additional information that accompanies more harmonic settings of music. In addition, melodies were played on a digital synthesizer under computer control to eliminate expressive deviations in note length and tempo.

6.3.2 Method

6.3.2.1 Subjects

The subjects consisted of 23 undergraduate students and graduate students drawn from the psychology courses at the University of Stirling. Undergraduate subjects were given credit for participation, partially fulfilling a degree requirement.

6.3.2.2 Equipment

Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a Yamaha DX-7 Synthesizer using its standard *Harp-Flute* patch. The synthesizer's output was synchronised with a NeXT TurboColor computer by means of a standard Midi connection. Subjects provided their responses by dragging a slider knob to a desired position by use of the computer's graphical user interface (GUI environment) and mouse as input device.

6.3.2.3 Procedure

Subjects heard an entire Gaelic folktune chosen from the complete set of tunes (see Section 6.3.2.4.1), computer synthesized and played through the NeXT computer's speaker. After hearing each tune, subjects moved a slider to indicate how happy or sad they felt the entire piece was (from *Happy* to *Sad*: see Figure 6.1). Subjects were not allowed to move on to the next trial until they had moved the slider at least once. If they did not move the slider, the program provided a message indicating that they must move the slider before continuing.

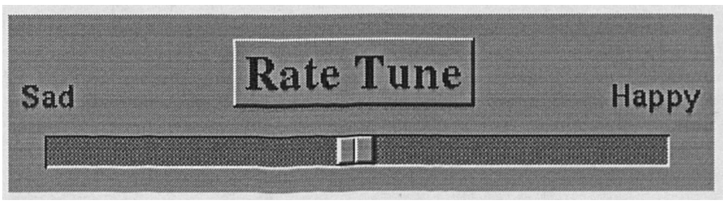


Figure 6.1. GUI slider for indicating mood rating after hearing entire Gaelic folktune.

6.3.2.4 Stimuli

6.3.2.4.1 Music

The stimuli consisted of the melodies line from thirty-one traditional Gaelic folk songs (MacPharlain, 1908) (see **Appendix A** for complete transcription of tunes), that would have originally been sung by a single, unaccompanied voice. These pieces varied in tempo and metre. The only interpretive elements were the addition of accents at the beginning of every bar, to provide a standard indication of pulse. All notes were produced with constant volumes within each tune and between tunes.

6.3.2.4.2 Instrument

The tunes were synthesized through the NeXT DSP chip, using a complex waveform with attack and decay envelopes, standard on the NeXT by the name of *Pluck*. This patch simulates the sound of a lute, guitar, or other stringed-instrument plucked string.

6.3.3 Results

The individual subject ratings were changed to rankings for each of the thirty-one separate pieces. Kendell's Coefficient of Concordance was computed for the rankings and yields significant agreement between the different subjects (W = 0.415; χ^2 = 286.186; $p < .001$). Several of the pieces were reliably judged as happy, and several as sad, indicating that music in this very simplified fashion could communicate affect. Table 6.1 shows the mean ratings, rankings, and standard deviations for all of the 31 gaelic folk tunes. One-sample *t*-test scores for each tune's ratings and their corresponding significance levels are shown in the rightmost columns. The table is ordered with the tunes rated furthest to the *Sad* pole of the continuum first, ascending in rating to those rated furthest to the *Happy* pole last.

Tune	Mean Rate	Mean Rank	SD Rate	SD Rank	*t* Score	*p* Level
1	-0.34	27.37	0.29	4.09	-5.642	*.001*
16	-0.27	24.74	0.35	7.06	-3.726	*.01*
30	-0.26	24.83	0.27	6.45	-4.645	*.001*
18	-0.21	23.74	0.27	5.86	-3.851	*.001*
2	-0.20	23.30	0.30	5.96	-3.214	*.01*
6	-0.18	22.11	0.31	5.98	-2.810	*.05*
22	-0.14	22.30	0.29	5.48	-2.271	*.01*
5	-0.07	20.37	0.36	8.29	-0.939	—
31	-0.06	19.72	0.28	6.88	-1.056	—
14	-0.02	17.85	0.30	7.00	-0.371	—
3	0.00	19.72	0.28	8.13	0.015	—
10	0.01	17.07	0.28	7.03	0.184	—
24	0.01	18.37	0.44	9.88	0.150	—
4	0.06	18.57	0.34	7.97	0.877	—
25	0.11	15.65	0.22	5.97	2.285	*.05*
11	0.11	16.48	0.31	8.07	1.729	—
12	0.13	15.54	0.22	6.88	2.869	*.01*
8	0.16	14.17	0.33	8.06	2.335	*.05*
29	0.20	13.48	0.32	7.33	2.907	*.01*
15	0.20	12.28	0.28	6.41	3.437	*.01*
13	0.20	15.24	0.31	7.84	3.120	*.01*
19	0.23	11.85	0.31	7.86	3.535	*.01*
7	0.30	10.89	0.31	7.26	4.538	*.001*
20	0.32	10.89	0.25	5.47	5.987	*.001*
26	0.34	9.11	0.26	4.33	6.333	*.001*
23	0.36	8.41	0.28	6.23	6.246	*.001*
28	0.38	8.83	0.34	7.42	5.366	*.001*
17	0.39	8.89	0.33	7.32	5.598	*.001*
27	0.40	9.11	0.33	8.03	5.742	*.001*
21	0.41	8.65	0.36	7.76	5.372	*.001*
9	0.43	6.91	0.29	4.80	7.203	*.001*

Table 6.1. Means and standard deviations of *Happy-Sad* ratings and rankings for all 31 gaelic folk tunes. The table is sorted with the tunes rated furthest to the *Sad* extreme first, neutral tunes in the middle, and those rated furthest to the *Happy* extreme last.

6.3.4 Discussion

The results provide evidence that these simple pieces do communicate information that is correlated with the *Happy/Sad* scale, and that subjects can reliably use that scale. Subjects then can regularly form transpositions from the affective qualities onto the experimenter imposed scale of *Happy/Sad.* It is difficult to conclude at any point that Piece A is actually happier than Piece B. Subjects were requested to draw their distinctions along this imposed scale; the use of this scale is another method of indirectly examining internal representations that correlate well with the experimental stimulus. This caution against actually attributing the qualities from the experimental scales as actual qualities of the stimulus will be important in later experiments, where several scales of measurement will be provided for stimulus evaluation.

Even with this caveat, the fundamental nature of the *Happy/Sad* scale suggests that the four tunes selected with the highest *Happy* ratings *are* actually happier in nature to the four with the lowest ratings. In the Osgood et al (1957) study, *evaluation* was suggested as one of the primary factors in establishing the semantic location of any word. The affective scale of *Happy/Sad* is basically a representation of evaluation—measurements on this scale may directly measure this one underlying psychological dimension.

6.4 Experiment Eleven: Melodic contributions to affect

6.4.1 Aims and Introduction

The ordering of tunes in Table 6.1 can be used to create a subset of tunes with the most extreme ratings for use in further experiments. Experiment Eleven took that subset and continued to examine alignments made between tunes and affective scales. That is, the stimulus songs were limited to those tunes that produced the most extreme *happy* ratings, the most extreme *sad* ratings. The first and last four tunes on this table are the extremes of the stimulus set both in terms of rating and ranking ordering.

This experiment began to explore the way in which a limited part of a melody can communicate affect, with the amount of music presented ranging along a continuum from entire melody to single note. It would certainly be unwieldy to examine all subsections of a melody for affective changes. Therefore, investigation of affect may begin with an examination of the melody's *extremities*: the beginning and ending groups of notes.

Experiment Eleven asked subjects to rate a subsection of each of those melodies drawn from Experiment Ten. Listeners were presented with sections of a limited number of beats, starting at the beginning or concluding on the final melodic note. At the conclusion of the passage, subjects are asked to rate the section on separate scales of happiness and sadness. It was assumed that the melody may have components of both moods expressed by the adjective pair; that is, that a melody could be both happy and sad in different parts. Subjects were therefore requested to rate the stimuli in terms of happy *and* sad, rather than using a single scale for indicating the stimulus's placement along a hypothesised single *Happy/Sad* continuum.

There were several aims to the experiment. It attempted to determine if short passages contained qualities that would be simultaneously rated on both ends of the affective continuum as outlined above. It also provided information on how many beats of music were necessary to produce an affective response in line with the piece's overall affective rating. Analysis of the subjects' ratings could also provide clues as to which particular aspects of the music produce affective ratings.

6.4.2 Method

6.4.2.1 Subjects

The subjects consisted of 14 undergraduate students and graduate students drawn from the psychology courses at the University of Stirling. Undergraduate subjects were given credit for participation, partially fulfilling a degree requirement. None of these subjects participated in Experiment Ten.

6.4.2.2 Equipment

The same equipment was used as in Experiment Ten.

6.4.2.3 Procedure

Subjects heard a portion of a gaelic folktune, and were asked to rate it on two sliding scales: how happy they felt the happiest part of the melody was, and how sad they evaluated the saddest part of the melody to be. The responses were recorded by the use of two independent GUI interface sliders (see Figure 6.2). The tunes were presented to each listener in random order. The listener had the opportunity to request the experimenter to replay the melody if they wished to hear it again before making their ratings. After rating the tune on both scales and pressing the "Set Rating" button, the next melody began automatically. After the presentation of all melody fragments, a "Thank You" message appeared and the subject was debriefed.

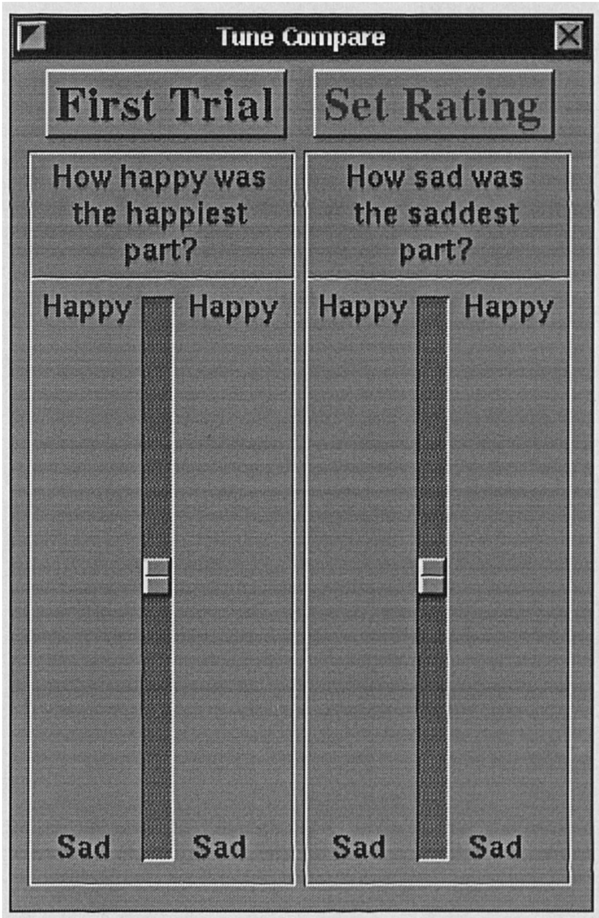


Figure 6.2. GUI sliders for indicating ratings for melody subsection.

6.4.2.4 Stimuli

6.4.2.4.1 Music Stimuli

Eight pieces were selected from the original thirty-one unaccompanied Gaelic folk tunes (*Tunes:* 1, 9, 16, 17, 18, 21, 27, & 30). The subsets of melodies can be divided into *Happy* and *Sad Subsets*, and designated with a subscript to indicate their listing position in Table 6.1 for reference. These subsets are given below:

Subset	Tune Number	Notation
Happy	9, 21, 27, & 17	H_1, H_2, H_3, H_4
Sad	1, 16, 30, & 18	S_1, S_2, S_3, S_4

Table 6.2. Tunes in Experiment Eleven stimulus subsets and notation shortcuts for tunes numbers. Notation order corresponds to tune position in column 2 (*e.g.*, H_1 corresponds to Tune 9, H_2 corresponds to Tune 21, etc.).

The tunes were presented in segments of varying beat length. The tunes were presented in segments of 2, 4, 6, 8, and 16 beats starting on the first note, and 2, 4, 6, 8, and 16 beats ending on the last note of the tune. See **Appendix A** for the transcriptions of all of the stimuli melodies (Tunes 1, 9, 16, 17, 18, 21, 27, 30). When referring to fragments from the different tunes, the term *beginning fragment* will designate those fragment composed of *n* beats from the beginning of the piece; the term *ending fragment* will be used for those fragments of *n* beats concluding on the tune's final note.

6.4.2.4.2 Instrument

The tunes were synthesized through the NeXT DSP chip, using a complex waveform with attack and decay envelopes, standard on the NeXT by the name of *Pluck*. This patch simulates the sound of a lute, guitar, or other stringed-instrument plucked string.

6.4.3 Results

Separate 3-way ANCOVAs were run on each of the two rating scales, measures of the happiest point and saddest point in the musical segment (hereafter referred to as *happy rating* and *sad rating*, respectively). These ANCOVAs were 2 x 2 x 5 designs, examining the independent factors of the tune's original rating category (Factor *Subset*: *Happy* or *Sad*, see Table 6.2), the part of the tune the beats were taken from (Factor *Position*: *Start* or *End*), and the number of beats played (Factor *Beats: 2, 4, 6, 8,* or *16*). The ANCOVA of one scale was performed with the other scale as a changing covariate. The first ANCOVA examined happy ratings with the sad ratings as covariate, and reveals significant main effects for Subset ($F(1, 1018) = 278.67, p < .001$) and Beats ($F(4, 1018) = 31.99, p < .001$). A two-way interaction between Subset and Beats is also present ($F(4, 1018) = 3.75, p < .01$), and plotted below in Figure 6.3. Post-hoc comparisons were performed on the main effect of Beats for the happy ratings, and reveals significant differences between ratings obtained for 2-note segments and all other segments (Scheffé Tests, all $p < .0001$).

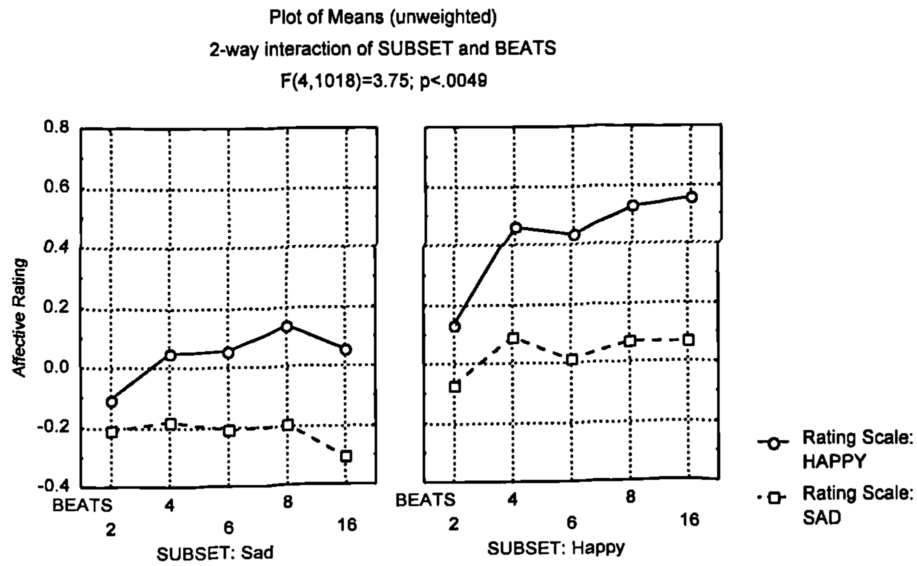


Figure 6.3. Two-way interaction of the Subset and Beat factors on the *Happy Ratings*, plotting with the covariate of *Sad Ratings*.

The second ANCOVA examined sad ratings with the happy ratings as covariate, and reveals a significant main effect for Subset ($F(1, 1018) = 50.74, p < .001$). No other main effects or interactions are observed in this ANCOVA.

Individual cells from the design were examined to determine if the happy rating or sad rating was significantly different from neutral, providing more information about a particular segment of a tune. This was performed to determine if an individual segment could be simultaneously rated as happy *and* sad. The 16-note segment of Tune 1 (S_1) was simultaneously rated happy on the happy rating ($t(12) = 3.033, p < .05$) and sad on the sad scale ($t(12) = -4.226, p < .01$). Figure 6.4 plots the happy and sad ratings for all of Tune 1's starting fragments together in a bar chart, and demonstrates the bi-directional rating pattern. In contrast, Figure 6.5 plots the happy and sad ratings for the last ending fragments of Tune 1, which is more indicative of the rating patterns found in the other tunes.

Figure 6.4–Figure 6.11 plot the means and standard deviations for each of the melody fragment ratings. In each of the plots the happy and sad ratings for the melody fragments composed of *n* beats from the beginning of the melody are presented on the left; the right portion of the graph plots the same information for fragments concluding on the last note of the melody. These ratings are plotted as the bars, with the scale of affective rating on the left ordinate. Line graphs of the happy rating and sad rating standard deviations are also plotted on the same graphs, with the ordinate for these measures on the right.

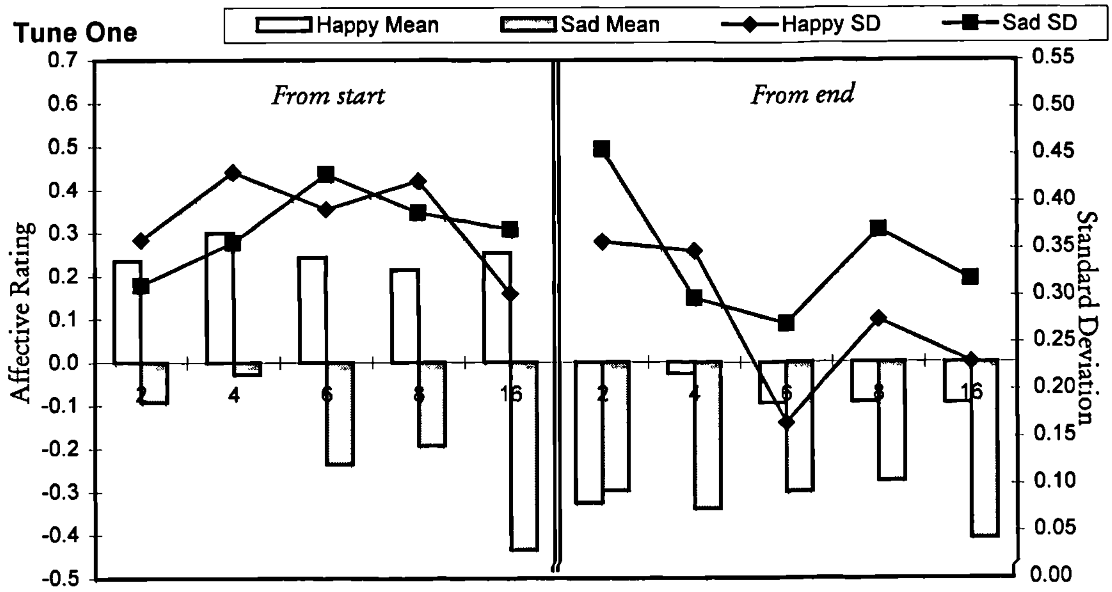


Figure 6.4. Happy and sad rating means and standard deviations for each of Tune One's (Sad_1) melody fragments.

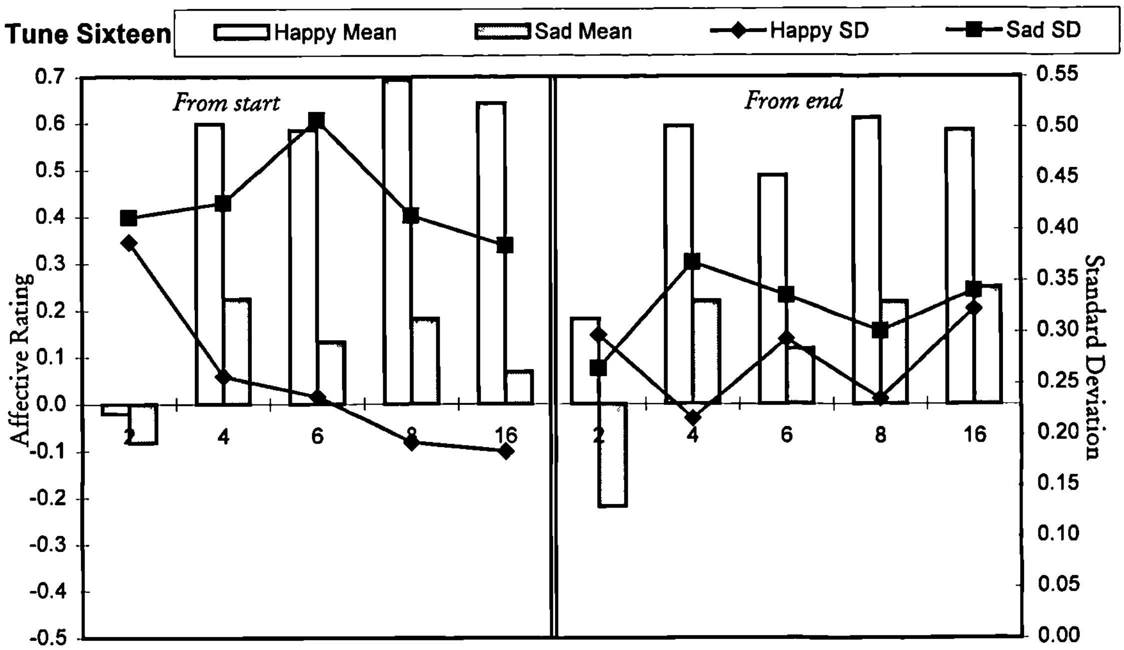


Figure 6.5. Happy and sad rating means and standard deviations for each of Tune Sixteen's (Sad_2) melody fragments.

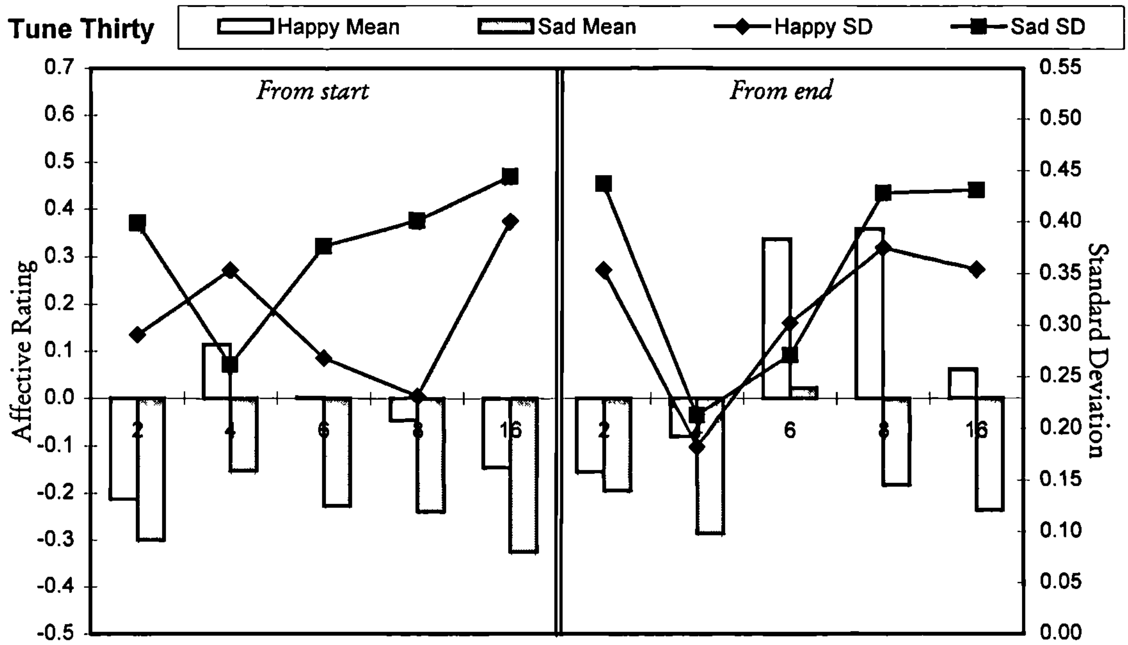


Figure 6.6. Happy and sad rating means and standard deviations for each of Tune Thirty's (Sad_3) melody fragments.

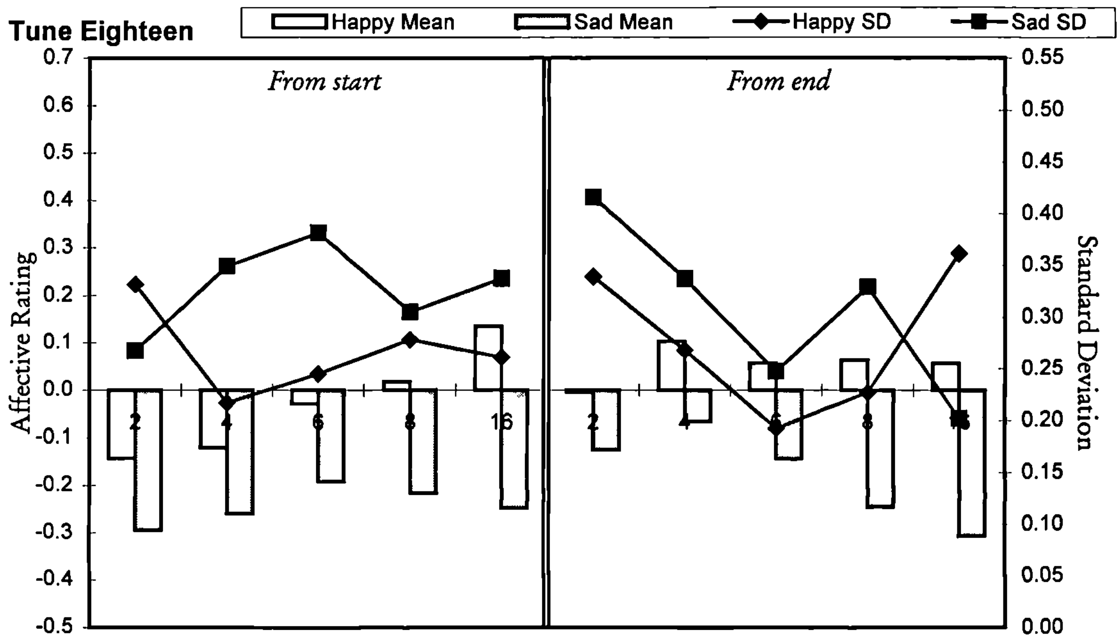


Figure 6.7. Happy and sad rating means and standard deviations for each of Tune Eighteen's (Sad_4) melody fragments.

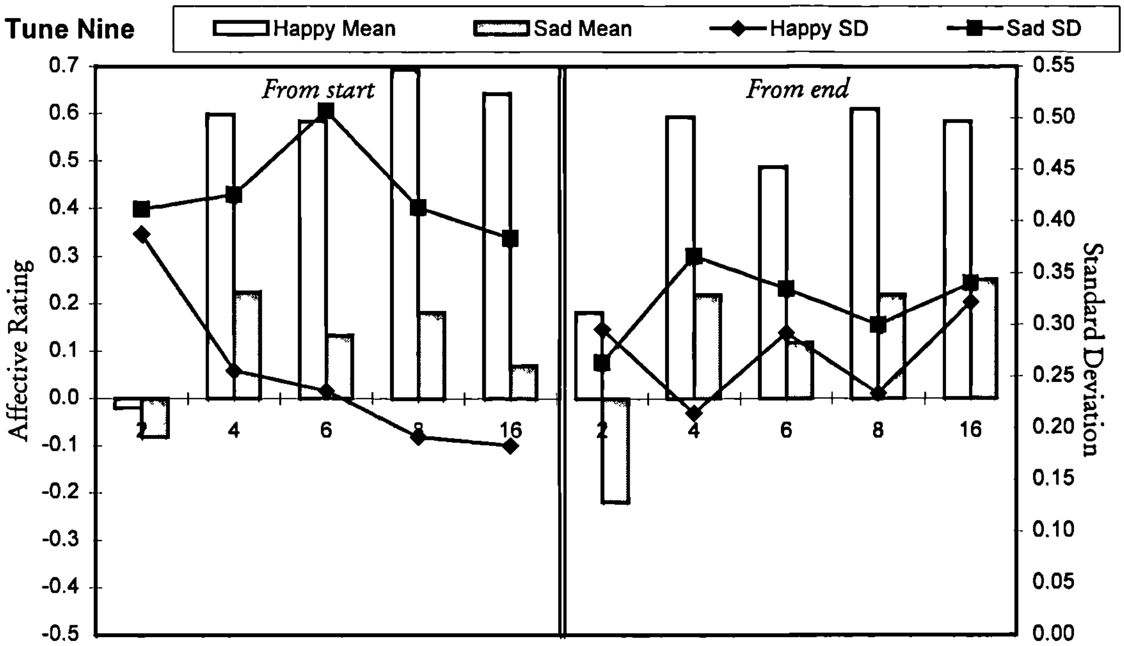


Figure 6.8. Happy and sad rating means and standard deviations for each of Tune Nine's ($Happy_1$) melody fragments.

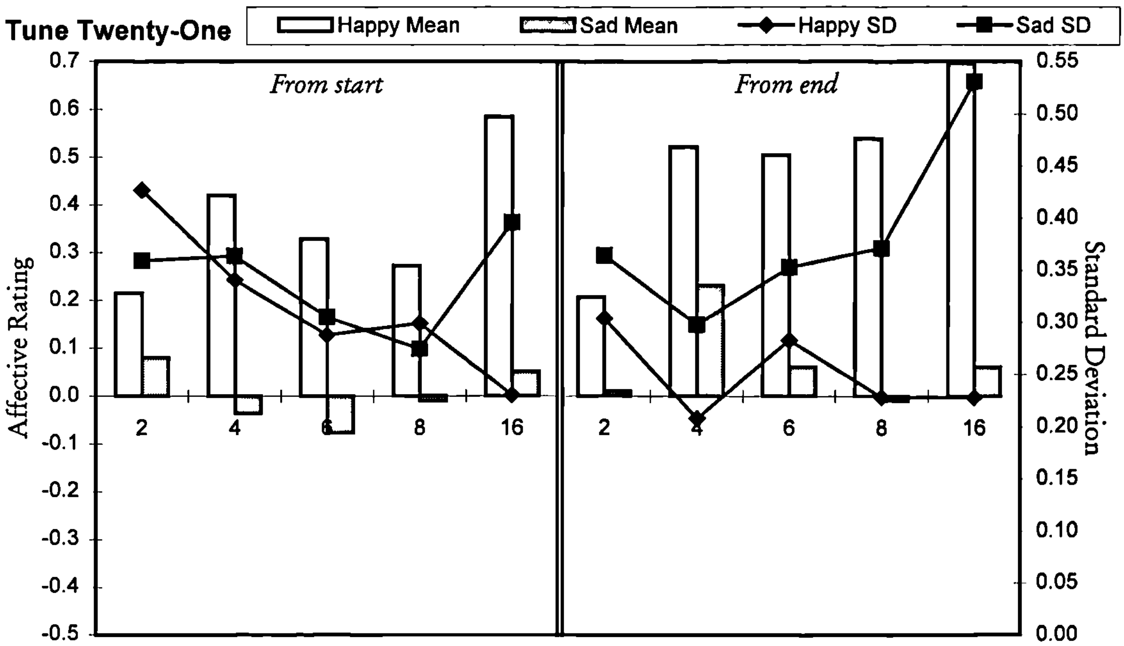


Figure 6.9. Happy and sad rating means and standard deviations for each of Tune Twenty-One's ($Happy_2$) melody fragments.

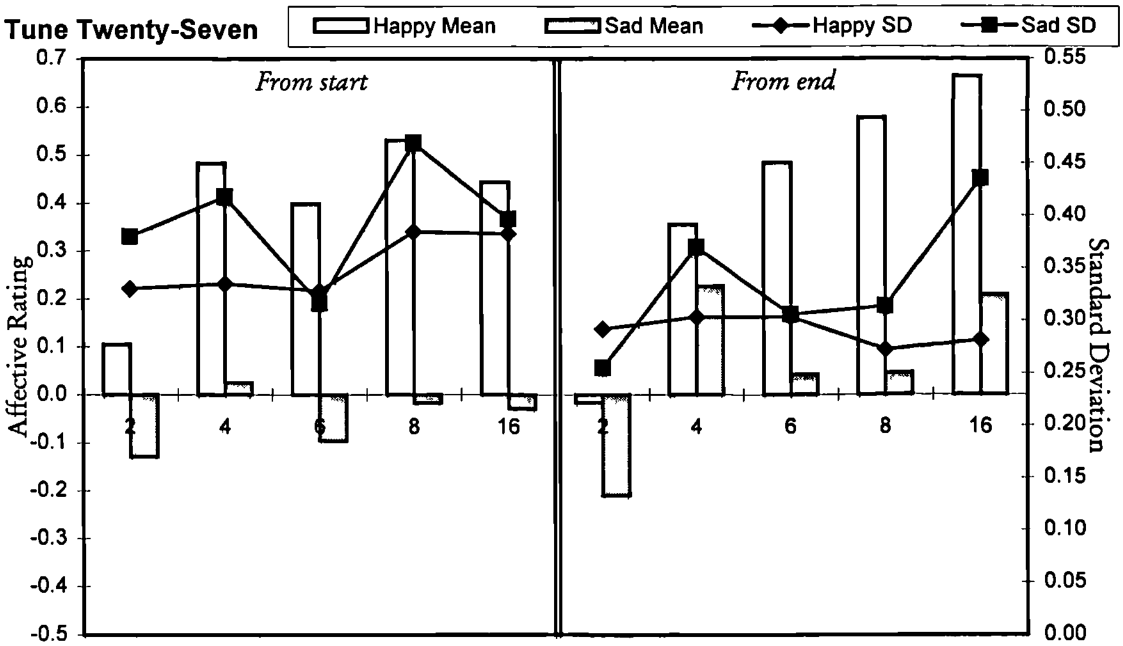


Figure 6.10. Happy and sad rating means and standard deviations for each of Tune Twenty-Seven's ($Happy_3$) melody fragments.

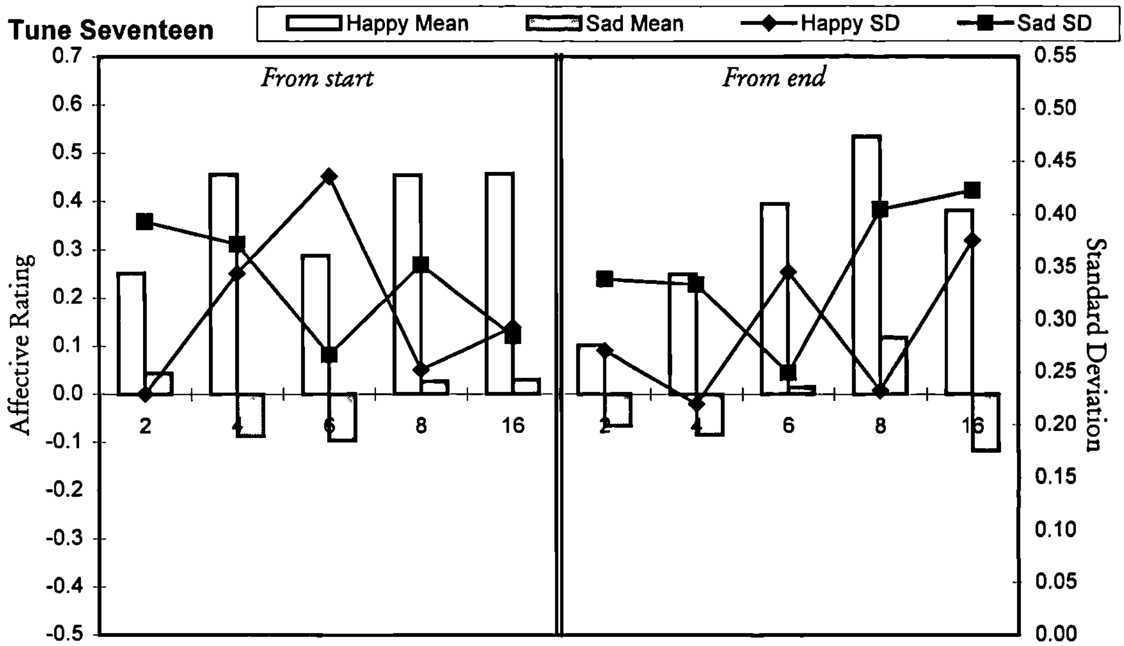


Figure 6.11. Happy and sad rating means and standard deviations for each of Tune Seventeen's ($Happy_4$) melody fragments.

6.4.4 Discussion

The most notable feature of the results is the number of beats it takes to establish the happy rating as significantly different from neutral, coupled with the lack of a similar result for the sad ratings. It takes four beats for the happy tunes to be established as happy. Post-hoc tests run on the separate happy segments reveal significant differences between the two beat segments and all other segments, but not between any other segment groups. This pattern of response suggests that the establishment of a happy rating is similar to a step function, with the step placed somewhere between two and four beats. This jump in rating is present for the happy tunes *and* sad tunes, indicating that both have qualities that can be rated as happy by subjects, and that the pattern occurs independently of the tune's overall affective rating.

Although the *Happy* and *Sad* rating scales had identical poles (ranging from *Happy* to *Sad*) and graphical layout, subjects were requested to use these scales to rate different characteristics of the tune stimuli. One of the aims of the experiment was to see if subjects would evaluate a piece of music as having both *happy* and *sad* qualities. In fact, the ratings for the beginning fragments of Tune One do reveal that such a pattern of response is possible. All of the *happy* ratings are on the happy side of the scale, while all of the *sad* ratings are on the sad side of the scale, with significant differences found between these ratings. This result may indicate affective ambiguity in the stimuli, or that subjects were matching of constructs with different affective SoCs. For example, it is possible that some quality of a fragment's *rhythm* is rated as happy while the quality of a fragment's *tonality* is rated as sad.

The ending fragments of Tune One are all rated on the sad side of the affective scales, indicating that the affective qualities of the tune are less ambiguous, and are represented in both scales. It is not unexpected that only one passage demonstrated simultaneous happy and sad ratings; rather, it is fortunate that to find even one tune in a stimulus set of simple tunes could reveal such a pattern. The relatively simple melodic structures of the tunes would likely lend themselves to unidirectional affective ratings, while more complicated melodies would be more likely to reveal more ambivalent affective ratings. The addition of polyphony in the tunes, making the stimuli more complex, would likely increase the variability of response for a single passage, in the same fashion that increasing stimulus complexity in Chapter Three increased the variance in cross-modal matching performance.

An overall decrease in standard deviation as a function of increasing number of beats for affective ratings is not observed. This suggests that subjects do not necessarily have greater concordance for affective content as a function of increased musical content. Some individual tunes do have standard deviations which decrease with increasing segment length, such as Tune Sixteen (S_2). Once again, the greatest change occurs between two and four beats. An overall effect of decreasing standard deviations with time is not found for all of the passages. Although the tune in its totality may produce uniform affective ratings by subjects, the affective ambiguity in a piece does not always decrease in a linear manner with more beats.

The results of Experiment Eleven demonstrate that an overall quality of sadness can be communicated by a tune, but this affect is established over a longer time period than the quality of happiness. Thus, it may be said that happiness in monophonic melodies can be communicated quicker than sadness. The relatively short length of the musical fragments that *can* establish high affective agreement across subjects suggests that evaluation or perception of simple musical constructs may account for a substantial part of the happy ratings. Longer passages will have higher-level qualities, such as a more clearly defined tonal centre, similarity to music previously known to the listener, and overall melodic structure and contour in addition to evaluation of the lower-level qualities. However, the evaluation of short passages is more likely to be formed on the basis of these lower-level qualities, such as intervallic leap size, intervallic leap direction, and syncopation.

If simple constructs help establish affective ratings, then the ratings of individual fragments compared to each other may provide some suggestions as to which constructs are important. The nature of such comparisons is necessarily speculative.

One example of a simple melodic construct potentially altering rating is the 2-beat fragment from Tune One. In this tune, the 2-beat segment consists only of a repeated note (B_3, dotted crotchet followed by two semi-quavers), while the four note segment adds only a two notes, but creates a syncopated rhythm. Subjects rated the repeated note as being sad on both available scales (and all but two subjects (77%) used the sad portion of the happy scale in their rating). Increasing the ending fragment from 2 to 4 beats of Tune One includes the introduction of four notes, and an overall syncopated rhythm. The minor tonal centre of the phrase (B-minor) may be enough to keep the sadness rating high, the syncopated rhythmic aspect of the phrase (dotted-quaver—semi-quaver, dotted-quaver—semi-quaver, minim) may partially account for the difference between happiness ratings. A paired *t*-test does reveal a significant difference between the means of the 2-beat and 4-beat ending fragment happiness ratings

($t(12) = -2.161$, $p < .05$). Additional examples of syncopated rhythms possibly contributing to happier ratings are the two-beat beginning fragment of Tune Twenty One and 2-beat ending fragment of Tune Nine.

The scale mode of the tunes may play a rather more obvious part in the affective ratings. All of the tunes are mostly pentatonic in nature, in either a minor or major mode. Establishment of a strong major mode, such as the first four beats of Tune 9, generates higher happiness ratings. In this tune, the initial two-beat beginning fragment consists of four quavers repeating an F_4. The only information inherent in this is the timbre of the instrument, overall rhythmic tempo, and the absolute pitch level of the note. Accordingly, the rating on both scales is close to neutral. Either the rhythmic repetition of the note, the higher frequency of this note, or a combination are sufficient to significantly change the happiness rating from the two-beat ending fragment in Tune One to the two-beat beginning fragment in Tune Nine ($t(12) = -2.449$; $p < .05$). Despite the different temporal positions of these fragments in their parent melody, it is still valid to compare their ratings as the fragments are too short to be considered as complete musical phrases.

When the two-beat beginning fragment of Tune Nine is increased to four-beats a major tonality is clearly defined with the descending pattern of the octave, major seventh, major sixth, and perfect fifth of an F-Major scale. Additionally, a syncopated rhythm is present (dotted-quaver—semi-quaver) by beat three. These melodic aspects significantly increase the happiness ratings along both scales (Happy:

$t(12) = -5.213; p < .001$; Sad: $t(12) = -2.480; p < .05$).

Another example of major modes affecting *Happiness* rating is the ending eight-beat fragment of Tune Sixteen. The increase of six beats to eight beats adds a set of triplets that create a major tonality in the fragment not present in the shorter fragments. In this fragment, the introduction of these two beats significantly increases happiness ratings along both scales (Happy: $t(12) = -4.257; p < .01$; Sad: $t(12) = -3.146; p < .01$).

A third melodic component is present in many examples that may contribute dramatically to *happiness* ratings: a large melodic leap. In Tune Thirty, two such examples are present, with both ascending and descending melodic leaps potentially altering the ratings. The change in both *Happy* and *Sad* scales found when comparing the 4-note ending fragment with the 6-note ending fragment is significantly more happy in rating (Happy: $t(12) = -4.112; p < .01$; Sad: $t(12) = -3.413; p < .01$). The leap here is a *descending* leap of an octave; however, the ascending leap present in the 4-note beginning fragment increases happiness ratings along both scales when compared to the 2-note beginning fragment. Several other instances of leaps increasing happiness ratings are present throughout the eight tunes.

6.5 Experiment Twelve: Multiple Categorical Evaluation of Gaelic Folk Tunes

6.5.1 Aims and Introduction

Experiments Ten and Eleven both examine whether affective content could be reliably communicated to listeners through music. The musical stimuli used in the experiments were more basic than recorded passages of music, as they were computer produced without human performance characteristics. The musical and auditory dimensions that were present to communicate to listeners were reduced to intervallic and temporal properties of the individual melodic notes.

Experiment Twelve pushed the paradigm from the earlier experiments further by examining the types of information that can be conveyed with simple musical stimuli in greater detail. To accomplish this, a stimulus rating task was used that required subjects to choose between two adjectives that designated extremes of evaluative scales. For example, the adjective pair of "Happy" and "Sad" represent the extremes of the affective scale used in Experiments Ten and Eleven. Other scales were gender ("Male" and "Female"), width ("Thin" and "Thick"), etc.

The task is essentially a variant on the *Semantic Differential* experiments of Osgood et al (1957). It differs in that it does not allow subjects to indicate the magnitude of their ratings; rather, it only forces a choice for one pole of the scale. This simplification of the rating task *eliminates the option of a neutral rating*. Subsequently, small biases may reveal sample trends that might not otherwise be observable. Analyses of chosen adjectives can reveal if there are preferences for one adjective in a pair regardless of the presented stimuli.

If it is found that adjectives are chosen in an unbiased manner for the overall stimulus set, but have a significant inter-subject concordance pattern for a single tune, it can be concluded that some qualities correlated with the adjective scales are communicated by the music. The strong results reported from earlier experiments using the same technique with recorded musical passages (Watt et al, *in preparation*) suggests that relatively small sample sizes can be used to examine such agreements effectively. In analysis, strong sample concordance with significance levels of p = .002 can be demonstrated with an N of 10.

6.5.2 Method

6.5.2.1 Subjects

The subjects consisted of 10 undergraduate students, and graduate students, five male and five female, drawn from the psychology courses at the University of Stirling. Undergraduate subjects were given credit for participation, partially fulfilling a degree requirement. None of these subjects participated in Experiments Ten or Eleven.

6.5.2.2 *Equipment*

Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a Yamaha CBX-T3 Synthesizer using a digitally sampled acoustic grand piano patch. The synthesizer's output was synchronised with a NeXT TurboColor computer by means of a standard Midi connection. Subjects provided their responses using a specially designed GUI interface to enter their choices.

6.5.2.3 *Procedure*

Subjects heard the entirety of a Gaelic folktune, and were asked to choose one of the two adjectives in each of the displayed pairs: *Joyful-Sad, Female-Male, Old-Young, Bright-Dark, Prickly-Smooth, Thick-Thin, Moist-Dry, Narrow-Wide, Angry-Pleased, Good-Evil, Leaden-Weightless, Sweet-Sour, Violet-Yellow, Day-Night* (see Figure 6.12). Subjects were able to enter responses in any order that they wished by means of clicking on the desired choice with the computer's mouse, or by using the left or right keyboard arrow keys to choose the left or right word of the currently selected adjective pair. The subject was required to rate each tune on all categories before proceeding on to the next tune. If they attempted to proceed without completing the ratings a warning message appeared and requested that they completed all ratings. The subject was allowed to listen to each tune as many times as they desired by clicking on the "Play Sound Again" button with the mouse, or simply by pressing the "0" key on the computer keypad. At the completion of rating the last tune, the subject was shown a "Thank You" message and debriefed as to the purpose of the experiment.

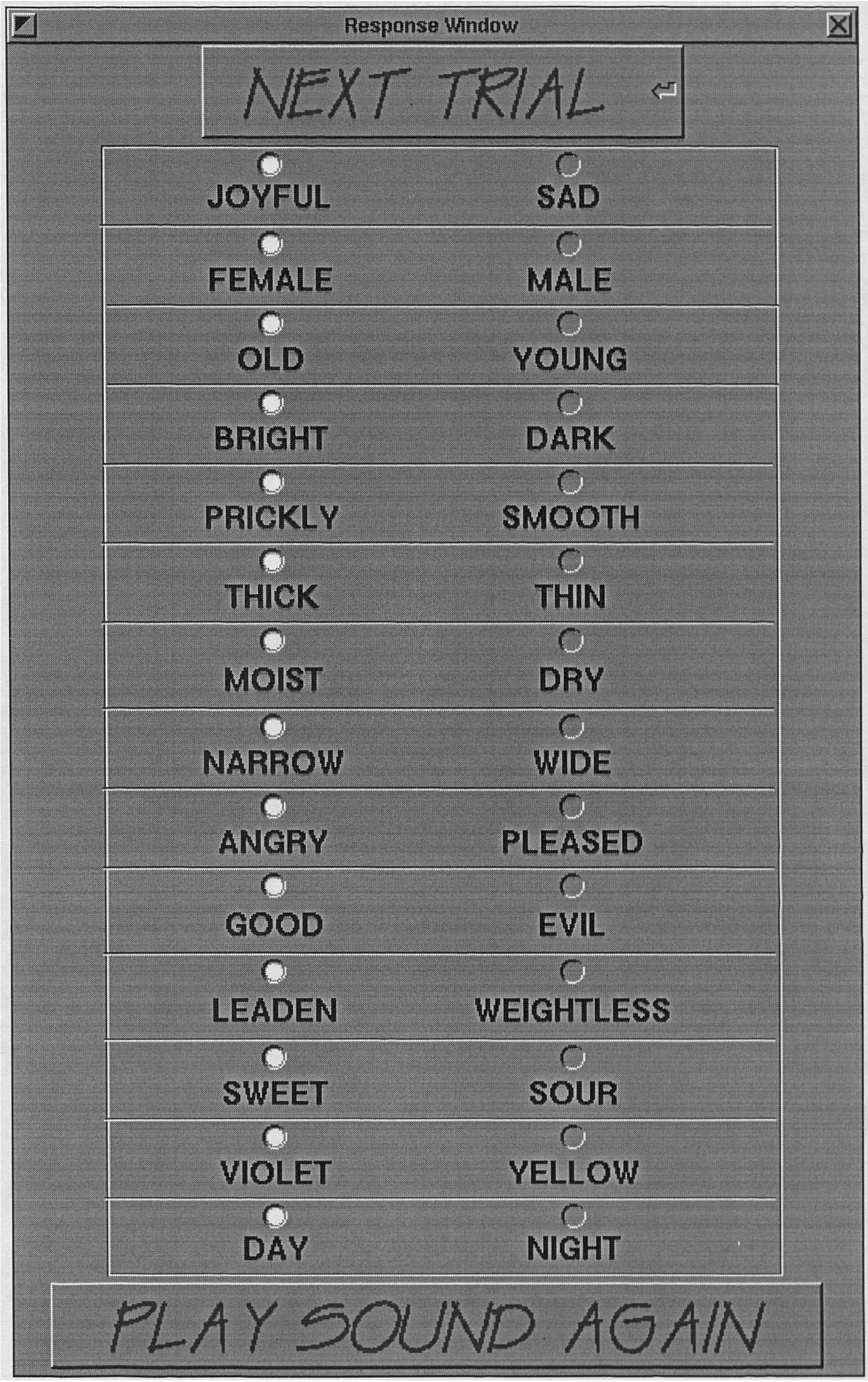


Figure 6.12. GUI interface for adjective pair ratings.

6.5.2.4 Stimuli

6.5.2.4.1 Music Stimuli

The pieces consisted of the complete set of thirty-one unaccompanied Gaelic folk tunes (see **Appendix A**).

6.5.2.4.2 Instrument

The tunes were synthesized through the Yamaha CBX-T3, using the sampled patch of an acoustic grand piano. The synthesizer was attached to an external set of amplified speakers, playing at an approximate volume of 78 dB.

6.5.3 Results

In order to determine inter-subject concordance, values were computed for the number of subjects who chose the same adjective for a tune. These values will be referred to as *concordance ratings.* To compute these, each adjective in a pair was assigned a value of 0 or 1. A score was computed by totaling up the adjective values chosen by all subjects for an individual tune. For example, "Good" was assigned a 0 and "Evil" was assigned a 1; a score of 0 indicates that all subjects selected "Good" and 10 indicates that all selected "Evil". These scores were transformed with the following formula, where *s* is the adjective pair score:

$$C = \frac{|s-5|}{10} + 0.5 \qquad (6.1)$$

This yields a value that ranges from 0.5 to 1.0, indicating percentage of concordance. Such a measure places no importance upon which particular adjective is chosen: high values demonstrate only that subjects agreed upon which adjective was appropriate for a stimulus. Using the binomial distribution (with $p = 0.5$ for any one trial), the probability of obtaining a concordance rating of 0.9 for an adjective pair by chance is $p = .022$, and obtaining 1.0 is $p = .002$.

Table 6.3 provides the concordance ratings for all adjective pairs and all tunes. Significant concordance ratings in the table are highlighted and emboldened for reference. The rows in the table are ordered by averaging across the concordance ratings, with those pieces having the highest average concordance ratings at the top. Thus, Tune 23 evoked the most concordance across all categories, followed by Tunes 27 and 9.

Audio	Joyful	Female	Old	Bright	Prickly	Thick	Moist	Narrow	Angry	Good	Leaden	Sweet	Violet	Day	Ave.
23	1	0.9	0.7	1	0.5	0.8	0.5	0.7	1	1	0.9	1	0.7	0.9	83%
27	0.8	0.8	0.9	1	0.6	0.7	0.5	0.5	0.8	1	0.7	0.9	0.8	1	79%
9	0.9	0.7	0.7	0.9	0.8	0.6	0.6	0.5	1	1	0.8	1	0.6	0.8	78%
1	1	0.7	0.8	0.9	0.8	0.6	0.6	0.8	0.9	0.5	0.9	0.8	0.6	0.8	76%
5	0.9	0.6	0.7	0.9	0.9	0.8	0.8	0.6	0.8	0.6	0.9	0.7	0.8	0.7	76%
17	0.9	0.6	0.6	0.8	0.8	0.6	0.6	0.7	0.8	1	0.6	0.9	0.8	0.8	75%
20	1	0.6	0.8	1	0.5	0.5	0.5	0.5	1	1	0.5	0.9	0.6	1	74%
18	0.9	0.6	0.7	0.7	0.9	0.6	0.7	0.7	0.9	0.6	0.8	0.7	0.9	0.6	74%
21	1	0.5	0.6	0.9	0.5	0.6	0.5	0.5	0.9	1	0.9	0.9	0.8	0.7	74%
30	0.8	0.7	0.7	0.8	0.7	0.6	0.6	0.8	0.8	0.6	0.9	0.9	0.8	0.6	74%
2	0.8	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.5	0.8	0.7	0.5	0.5	0.6	72%
16	0.8	0.6	0.8	0.8	0.7	0.7	0.7	0.8	0.6	0.8	0.8	0.6	0.6	0.8	72%
10	0.6	0.6	0.8	0.7	0.9	0.7	0.8	0.7	0.5	0.9	1	0.6	0.7	0.5	71%
4	0.6	0.5	0.7	0.6	0.7	0.8	0.6	0.9	0.6	0.9	0.9	0.8	0.6	0.6	70%
13	0.7	0.8	0.6	0.7	0.7	0.5	0.6	0.6	0.8	0.9	0.6	1	0.6	0.7	70%
26	1	0.6	0.6	0.7	0.7	0.7	0.6	0.8	0.9	0.9	0.6	0.7	0.5	0.5	70%
7	0.7	0.6	0.5	0.8	0.7	0.6	0.6	0.6	0.7	1	0.7	0.8	0.6	0.8	69%
22	0.7	0.7	0.7	0.8	0.7	0.8	0.5	0.9	0.6	0.8	0.6	0.7	0.6	0.5	69%
3	0.6	0.7	0.6	0.7	0.5	0.6	0.6	0.7	0.8	1	0.6	0.8	0.6	0.5	66%
24	0.8	0.7	0.5	0.6	0.8	0.6	0.6	0.7	0.7	0.8	0.8	0.7	0.5	0.5	66%
28	0.5	0.6	0.8	0.5	0.9	0.5	0.5	0.5	0.6	1	0.6	0.8	0.8	0.6	66%
19	0.6	0.6	0.7	0.8	0.6	0.6	0.5	0.7	0.6	0.9	0.5	0.5	0.8	0.6	64%
31	0.7	0.5	0.5	0.7	0.7	0.8	0.5	0.6	0.7	0.8	0.8	0.5	0.5	0.7	64%
12	0.5	0.7	0.7	0.5	0.6	0.5	0.8	0.5	0.8	0.9	0.6	0.7	0.5	0.6	64%
25	0.8	0.5	0.6	0.7	0.7	0.5	0.6	0.5	0.8	0.9	0.5	0.8	0.5	0.5	64%
29	0.7	0.5	0.7	0.7	0.6	0.7	0.6	0.6	0.5	0.8	0.6	0.7	0.5	0.6	63%
8	0.5	0.8	0.7	0.7	0.5	0.5	0.5	0.5	0.6	0.8	0.7	0.5	0.6	0.8	62%
6	0.7	0.5	0.8	0.5	0.6	0.7	0.5	0.6	0.6	0.8	0.7	0.6	0.5	0.5	61%
15	0.5	0.8	0.6	0.7	0.5	0.5	0.6	0.5	0.5	0.8	0.8	0.5	0.6	0.7	61%
11	0.6	0.6	0.5	0.5	0.6	0.6	0.5	0.6	0.5	0.7	0.5	0.8	0.7	0.8	61%
14	0.7	0.5	0.6	0.6	0.5	0.7	0.5	0.6	0.6	0.7	0.8	0.5	0.6	0.5	60%
	75%	64%	68%	74%	68%	64%	59%	65%	72%	85%	72%	74%	64%	67%	

Table 6.3. Concordance ratings for each tune across each of the adjective pairs.

Chi-square values were computed for each column, to determine if an individual adjective pair was used significantly more than chance, by using the following equation:

$$\chi^2 = \sum_{i=1}^{31} \left(\frac{score_i - p}{\sqrt{npq}} \right)^2 \qquad (6.2)$$

...where p and q are both 0.5. This produces a chi-square with 31 degrees of freedom that can be examined for significance. Table 6.4 shows the computed values and significance levels for the adjectives found to have highly significant chi-square values.

In order to determine that scales were being used in a meaningful manner (*i.e.*, subjects were not *always* choosing one of the adjectives over another), the total score for the adjective across all subjects and tunes was computed, and the probability determined from the binomial distribution for $p = .05$. Significant biases observed for the scales in Table 6.4 are indicated in the column titled "Bias" with the adjective from the pair predominantly selected by subjects. Table 6.5 shows the correlation coefficients computed for each of the adjective pair combinations.

Adjective	χ^2	*p-level*	Bias
Good/Evil	171.847	*.001*	Good
Joyful/Sad	110.559	*.001*	
Sweet/Sour	99.744	*.001*	Sweet
Bright/Dark	98.141	*.001*	
Angry/Pleased	92.533	*.001*	Pleased
Leaden/Weightless	84.922	*.001*	Leaden
Day/Night	62.891	*.001*	Day
Prickly/Smooth	60.487	*.01*	Smooth
Old/Young	52.476	*.01*	

Table 6.4. Correlations for all adjective pair scores.

In order to determine if subjects were consistently selecting only one adjective of a pair to represent all tunes, scores were summed across all tunes for each subject. This analysis revealed that three subjects always selected a single adjective for all tunes: one selected *Good*, another *Violet*, and another *Old* on every trial. Another subject chose the five adjectives of *Thick, Moist, Wide, Heavy,* and *Yellow* as indicative of all tunes.

Audio	Joyful	Female	Old	Bright	Prickly	Thick	Moist	Narrow	Angry	Good	Leaden	Sweet	Violet	Day
Joyful	—	.48	-.78	.91	.36	-.61	.56	.44	-.87	.80	-.75	.80	-.63	.72
Female		—	-.63	.68	.57	-.58	.49	.57	-.51	.44	-.45	.47	-.33	.64
Old			—	-.85	-.47	.57	-.58	-.51	.66	-.49	.68	-.63	.51	-.78
Bright				—	.53	-.73	.60	.65	-.80	.77	-.77	.75	-.60	.85
Prickly					—	-.39	.31	.39	-.42	.26	-.39	.28	-.49	.48
Thick						—	-.40	-.54	.55	-.47	.72	-.64	.48	-.64
Moist							—	.34	-.48	.41	-.51	.52	-.41	.57
Narrow								—	-.45	.41	-.51	.47	-.13	.50
Angry									—	-.84	.72	-.85	.64	-.58
Good										—	-.64	.76	-.57	.60
Leaden											—	-.81	.64	-.65
Sweet												—	-.62	.62
Violet													—	-.61
Day														—

Table 6.5. Correlations for all adjective pair scores.

A principle component analysis was performed on the scores obtained for each adjective and each tune. A score was computed from the number of times the second adjective of a pair was chosen. The analysis extracted two factors with eigenvalues in excess of 1.0. The first of these two factors accounts for more than 60% of the total variance (Factor 1: Eigenvalue = 8.671; 61.935% of total variance), while the second accounts for almost 10% (Factor 2: Eigenvalue = 1.280; 9.14% of total variance). Table 6.6 shows the factor loadings for each of these factors. Figure 6.13 graphically plots this data in two-dimensions after a varimax rotation has been performed on the two factors.

Variable	*Factor1*	*Factor2*
JOYFUL	* - 0.904	0.250
FEMALE	* - 0.703	- 0.510
OLD	* 0.839	0.156
BRIGHT	* 0.963	- 0.065
PRICKLY	- 0.551	- 0.478
THICK	* 0.760	0.149
MOIST	- 0.651	- 0.088
NARROW	- 0.622	- 0.445
ANGRY	* 0.868	- 0.294
GOOD	* - 0.787	0.372
LEADEN	* 0.852	- 0.158
SWEET	* - 0.856	0.311
VIOLET	0.698	- 0.318
DAY	* - 0.847	- 0.156
Expl. Var.	8.671	1.280

Table 6.6. Unrotated factor loadings for the two principle components in Experiment Twelve.

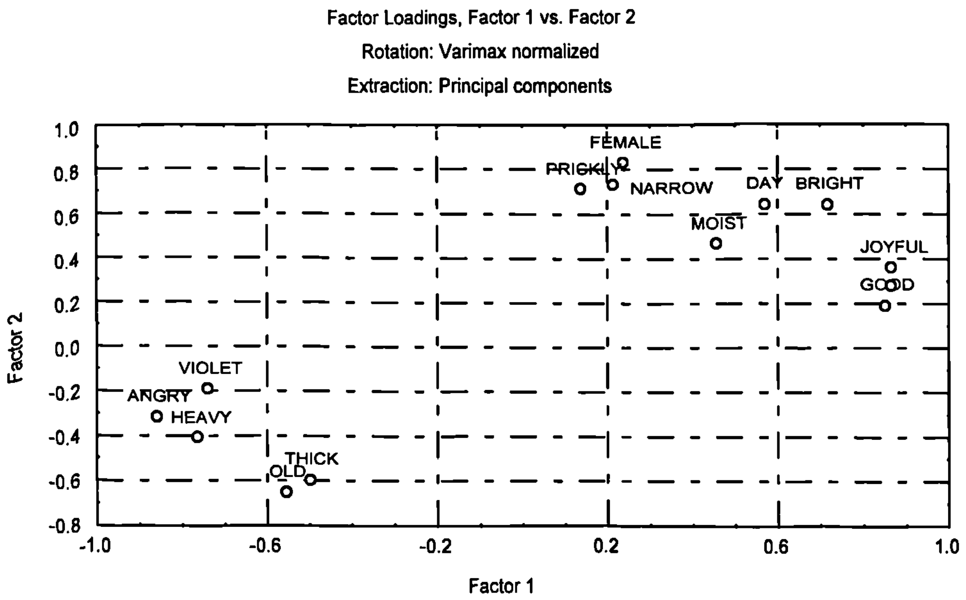


Figure 6.13. Factor loadings for each adjective pair after varimax rotation.

6.5.4 *Discussion*

The eight pieces that were selected from Experiment Ten for use in Experiment Eleven for their extremes in affective rating (Tunes 9, 21, 27, 17, 1, 16, 30, & 18; H_1–H_4 & S_1–S_4, respectively) were among the top 11 pieces from the average concordance ratings. This indicates that the pieces were among the less ambiguous pieces in categorical content as well as affective content.

Similar to the results of Watt et al (*in preparation*), several of the adjective categories with the highest concordance ratings are those adjectives that would be used to describe human personality characteristics. For example, *Good/Evil*, *Joyful/Sad*, and *Angry/Pleased* which were among the top five highest agreed-upon pairs. One of the remaining top five pairs also figured high in the Watt experiment, and might be used to describe human personality or taste sensation: *Sweet/Sour*. It is possible that subjects treated this adjective pair in both experiments as a personality characteristic. In terms of the SoC hypothesis, the associations formed by using *Sweet/Sour* as an affective SoC rather than a sensory scale results in more consistent matching. That is, the affective components communicated by the music may be closer in cognitive space to those of personality characteristics than gustatory characteristics.

These ratings cannot be used to form a definition of what a tune accurately communicates: total agreement of subjects to rate a piece as *Male* rather than *Female* does not mean that the piece *is* male. As subjects were forced to make a decision between ratings, it can only be said that subjects rated the piece as more masculine than feminine. The underlying factors that are being used to rate the pieces can only be conjectured from pairs contributing to each factor. Osgood et al (1957) suggested that three factors accounted for over 50% of the total variance in responses: evaluation, potency, and activity. The first factor in the present study appears to be an affective rating of "agreeability" or "pleasureability", which supports the existence of a primary evaluative factor in musical semantic differentials, as well as in language semantic differentials. Those elements that would closely correlate with this type of measure, such as *Joyful/Sad*, *Bright/Dark*, and *Sweet/Sour* each contribute greatly to the first factor. The second factor is more personality characteristically based, with the gender distinction *Male/Female* and possible human characteristic *Prickly/Smooth* contributing to the factor.

The agreements observed in Experiment Twelve are the result of either common scales of comparison aligned in a regular manner across subjects to the adjective pair scales, or the use of scales sufficiently similar across subjects to correlate well. Although these scales may be difficult for a subject to verbalise, the method of using the adjective pairs as a rating mechanism allows the experimenter to indirectly examine some of these subjects' internal representations of the music.

6.6 Experiment Thirteen: Evaluation of Gaelic Folk Tunes and Landscape Images

6.6.1 Aims and Introduction

Experiment Twelve demonstrated that monophonic music could reliably convey semantic information to people unfamiliar with the tunes or intended (lyrical) content. The two principal components extracted from the results might be all of the information that can be collected using a semantic differential method with simple tunes as stimuli. If so, such a limitation could be a product of using language SoCs for comparisons with music.

The earlier cross-modal chapters demonstrate that musical elements can be regularly matched with simple visual stimuli—subjects might also be able to match much more complex visual stimuli with complete tunes. Experiment Thirteen pushed the paradigm of a simple 2AFC cross-modal matching task even further than Experiment Twelve by asking subjects to match the Gaelic folktunes with photographs of varying landscapes. This experiment asked two questions: 1) can subjects match perceived qualities from simple Gaelic folktunes regularly to qualities in landscapes, and 2) if they can, are there a different number of factors used to make such matchings.

6.6.2 *Method*

6.6.3 *Subjects*

The subjects consisted of 10 undergraduate students, and graduate students, three male and seven female, drawn from the psychology courses at the University of Stirling. Undergraduate subjects were given credit for participation, partially fulfilling a degree requirement. None of these subjects participated in any of the preceding experiments.

6.6.3.1 *Equipment*

Audio stimuli were presented via a pair of Sony CD450 headphones, and produced by a Yamaha CBX-T3 Synthesizer using a digitally sampled acoustic grand piano patch. The synthesizer's output was synchronised with a NeXT TurboColor computer by means of a standard Midi connection. Subjects provided their responses using a specially designed GUI interface to enter their choices.

6.6.3.2 *Procedure*

Subjects heard the entirety of a Gaelic folktune, and were asked to choose between one of two pictures affixed to a page (see **Appendix B**). The subject's responses to that picture were entered into the computer using the same GUI interface used in

Experiment Twelve, with a changed set of labels for the picture pairs in place of the adjective pair (*e.g.*, the first adjective pair *Joyful/Sad* became *Picture 1/Picture 2*). There were a total of 14 picture pairs. To control partially for ordering effects, subjects began rating pictures on a different page for each trial, and alternated proceeding through the pictures in a forward and reverse serial sequence, led by the experimenter. At the conclusion of rating the last tune subjects were debriefed.

6.6.3.3 Stimuli

6.6.3.3.1 Music Stimuli

The pieces consisted of the complete set of thirty-one unaccompanied Gaelic folk tunes (see **Appendix A**).

6.6.3.3.2 Pictorial Stimuli

The pictures were taken from picture books of Scottish landscapes and city scenes, roughly balanced for size. The complete set of pictures is reproduced in **Appendix B**. In selecting the pictures, effort was given to selecting pictures varied in a subjective quality. For example, one picture was of a calm seascape, while another showed a rocky water scene with wavy whitewater.

6.6.3.3.3 Instrument

The tunes were synthesized through the Yamaha CBX-T3, using the sampled patch of an acoustic grand piano.

6.6.4 Results

The same types of analyses carried out on the results of Experiment Twelve were again employed for the results of Experiment Thirteen. Concordance ratings were computed for each adjective pair using Equation 6.1. Table 6.7 provides the concordance ratings for all picture pairs and all tunes. Significant concordance ratings in the table are highlighted and emboldened for reference. The rows in the table are ordered by averaging across the concordance ratings, with those pieces having the highest average concordance ratings at the top. Thus, Tune 1 evoked the most concordance across all categories, followed by Tunes 30 and 18. Table 6.8 shows the correlation coefficients for each of the fourteen stimuli picture pairs.

	Landscape Picture Pairs														
Audio	1–2	3–4	5–6	7–8	9–10	11–12	13–14	15–16	17–18	19–20	21–22	23–24	25–26	27–28	Ave.
1	0.9	0.6	0.6	0.8	0.7	0.7	1	0.9	1	1	0.8	1	0.8	0.8	83%
30	0.8	0.7	0.6	0.8	0.5	0.7	0.9	0.6	1	0.9	0.6	0.8	0.8	0.9	76%
18	0.6	0.5	0.5	0.6	0.7	0.7	0.9	0.7	0.9	1	0.8	0.8	0.9	0.9	75%
2	0.7	0.5	0.9	0.8	0.5	0.6	1	0.6	0.7	1	0.5	0.8	0.8	0.8	73%
16	0.7	0.6	0.7	0.7	0.6	0.8	0.7	0.6	0.9	0.7	0.7	1	0.6	0.8	72%
5	0.6	0.6	0.6	0.9	0.6	0.8	0.5	0.6	0.8	0.8	0.7	0.8	0.9	0.8	71%
22	0.5	0.6	0.5	0.6	0.7	0.7	0.6	0.7	0.9	0.9	0.6	0.7	0.9	1	71%
25	0.8	0.7	0.8	0.7	0.6	0.6	0.8	0.7	0.5	0.6	0.5	0.8	0.9	0.8	70%
31	0.8	0.5	0.7	0.7	0.7	0.8	0.9	0.5	0.7	0.6	0.5	0.7	0.8	0.7	69%
9	0.7	0.7	0.7	0.6	0.5	0.7	0.8	0.7	0.9	0.7	0.5	0.8	0.6	0.6	68%
21	0.7	0.9	0.6	0.8	0.8	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.8	68%
15	0.8	0.9	0.6	0.7	0.6	0.6	0.5	0.6	0.7	0.5	0.6	0.8	0.7	0.8	67%
28	0.5	0.8	0.7	0.7	0.5	0.9	0.6	0.6	0.7	0.6	0.8	0.8	0.5	0.7	67%
3	0.5	0.7	0.6	0.6	0.5	0.9	0.5	0.6	0.6	0.7	0.7	0.7	0.9	0.8	66%
27	0.8	0.7	0.5	0.6	0.7	0.7	0.7	0.5	0.6	0.5	0.5	1	0.5	1	66%
23	0.7	0.5	0.7	0.8	0.5	0.6	0.8	0.6	0.7	0.5	0.7	0.9	0.6	0.7	66%
17	0.9	0.9	0.7	0.5	0.5	0.7	0.7	0.6	0.7	0.6	0.5	0.8	0.5	0.6	66%
11	0.6	0.8	0.7	0.6	0.5	0.8	0.6	0.8	0.6	0.7	0.6	0.5	0.8	0.5	65%
19	0.7	0.5	0.7	0.7	0.6	0.6	0.6	0.7	0.5	0.7	0.6	0.6	0.8	0.8	65%
6	0.7	0.5	0.6	0.7	0.5	0.8	0.7	0.5	0.7	0.6	0.6	0.8	0.6	0.7	64%
8	0.8	0.7	0.6	0.5	0.7	0.5	0.6	0.6	0.5	0.7	0.7	0.8	0.6	0.7	64%
10	0.7	0.7	0.8	0.6	0.6	0.6	0.5	0.6	0.6	0.5	0.8	0.6	0.8	0.6	64%
12	0.6	0.6	0.7	0.7	0.5	0.6	0.6	0.8	0.5	0.7	0.7	0.6	0.7	0.7	64%
13	0.7	0.6	0.6	0.5	0.7	0.7	0.6	0.7	0.8	0.6	0.6	0.7	0.7	0.5	64%
26	0.7	0.7	0.5	0.7	0.5	0.8	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	64%
7	0.7	0.7	0.8	0.5	0.5	0.8	0.7	0.6	0.6	0.5	0.6	0.5	0.5	0.8	63%
29	0.8	0.9	0.5	0.7	0.5	0.5	0.6	0.7	0.7	0.5	0.6	0.5	0.7	0.6	63%
14	0.5	0.5	0.5	0.5	0.7	0.6	0.6	0.6	0.7	0.8	0.6	0.5	0.7	0.9	62%
4	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.7	0.8	0.8	0.7	0.6	0.6	0.5	61%
24	0.5	0.6	0.6	0.6	0.6	0.9	0.7	0.5	0.7	0.6	0.5	0.6	0.5	0.7	61%
20	0.7	0.6	0.6	0.6	0.6	0.7	0.5	0.5	0.8	0.5	0.8	0.5	0.5	0.6	61%
	69%	65%	64%	65%	59%	69%	68%	64%	71%	68%	63%	72%	69%	74%	

Table 6.7. Concordance ratings for each tune across each of the landscape picture pairs.

	Landscape Picture Pairs													
	1–2	3–4	5–6	7–8	9–10	11–12	13–14	15–16	17–18	19–20	21–22	23–24	25–26	27–28
1–2	—	.70	.14	–.42	–.43	.25	.73	–.52	.65	.61	–.28	–.62	.50	–.27
3–4		—	–.01	–.49	–.28	.01	.54	–.40	.60	.67	–.27	–.55	.41	–.34
5–6			—	–.16	–.11	.07	.41	.04	.13	.10	.15	–.27	.29	.12
7–8				—	–.06	.15	–.55	.37	–.49	–.58	.13	.61	–.41	.31
9–10					—	–.33	–.27	.26	–.25	–.15	–.09	.28	–.27	–.06
11–12						—	.30	–.02	.12	–.03	–.03	–.10	.03	.07
13–14							—	–.49	.68	.65	–.16	–.79	.53	–.32
15–16								—	–.49	–.49	.04	.40	–.13	.56
17–18									—	.77	–.35	–.75	.55	–.47
19–20										—	–.37	–.64	.63	–.38
21–22											—	.21	–.04	.02
23–24												—	–.65	.29
25–26													—	–.25
27–28														—

Table 6.8. Correlations for all landscape pair ratings across all tunes.

Table 6.9 shows the values computed with Equation 6.2 along with significance levels for the picture pairs found to have highly significant chi-square values. In order to determine that scales were being used in a meaningful manner (*i.e.*, subjects were not *always* choosing one of the adjectives over another), the total score for the adjective across all subjects and tunes was computed, and the probability determined from the binomial distribution (for $p = .05$). Significant biases observed for the scales in Table 6.9 are indicated in the column titled "Bias" with the picture from the pair predominantly selected by subjects. Table 6.9 shows the correlation coefficients computed for each of the adjective pair combinations.

Picture Pair	χ^2	*p-level*	Bias
27/28	90.130	*.001*	28
23/24	85.323	*.001*	
17/18	78.914	*.001*	17
19/20	68.499	*.001*	19
25/26	68.499	*.001*	25
13/14	65.695	*.001*	13
11/12	62.490	*.001*	11
1/2	59.285	*.01*	2

Table 6.9. Chi-square values, significance levels, and biases computed for each of the picture pairs. Values are sorted by chi-square. Bias determined by binomial distribution with p of 0.5 at a significance level of $p < .01$. Blank bias cells indicate that no significant bias for picture was found.

In order to determine if subjects were consistently selecting only one picture from a presented pair to represent all tunes, scores were summed across all tunes for each subject. This analysis revealed that one subject always selected a single picture for all tunes: Picture 19 from *19/20*, Picture 25 from *25/26*, and Picture 28 from *27/28*. All other subjects selected each picture to represent at least one tune.

A principal component analysis was performed on the scores for each tune and each adjective pair. A score was computed from the number of times the second picture from a presented pair was chosen. The analysis extracted four factors with eigenvalues in excess of 1.0. Factor One accounts for 42.80% of the total variance (Eigenvalue = 5.991) while Factor Two, Three and Four each account for approximately 30% of the remaining variance (Factor 2: Eigenvalue = 1.668; 11.91%

of total variance; Factor 3: Eigenvalue = 1.338; 9.56% of total variance; Factor 4: Eigenvalue = 1.182; 8.44% of total variance). Table 6.10 shows the factor loadings for each of these factors after a normalised varimax rotation has been performed in an attempt to clarify the structure of the factors in relation to the picture pairs.

Variable	*Factor1*	*Factor2*	*Factor3*	*Factor4*
1–2	0.664	0.445	0.225	-0.294
3–4	0.605	0.149	0.330	- 0.364
5–6	0.509	0.067	- 0.535	0.416
7–8	* - 0.711	0.300	- 0.017	0.245
9–10	- 0.179	* - 0.784	0.118	0.107
11–12	0.007	* 0.779	0.056	0.102
13–14	* 0.825	0.299	- 0.040	- 0.801
15–16	- 0.307	- 0.176	0.022	* 0.801
17–18	* 0.748	0.155	0.266	- 0.346
19–20	* 0.785	- 0.019	0.304	- 0.296
21–22	- 0.238	0.033	* - 0.875	- 0.133
23–24	* - 0.863	- 0.137	- 0.040	0.144
25–26	* 0.774	0.061	- 0.109	0.008
27–28	- 0.240	0.153	0.033	* 0.798
Expl. Var.	5.011	1.727	1.409	2.033

Table 6.10. Varimax rotated factor loadings for the four principle components in Experiment Thirteen.

Table 6.11 presents the tunes ordered by highest average concordance ratings; those on the left are ordered for adjectives and those on the right by picture pairs. The left-most column in each section indicates the tunes ranking on the affective *Happy/Sad* Scale recorded in Experiment Ten.

H/S	Audio	Adject	Agree	Pict	Agree	H/S	Audio	Adject	Agree	Pict	Agree
	7	15	69%	5	63%	H_4	17	26	75%	15	66%
	4	16	70%	2	61%		23	31	83%	16	66%
	13	17	70%	12	64%		3	12	66%	17	66%
	26	18	70%	7	64%	H_3	27	30	79%	18	66%
	10	19	71%	10	64%		28	11	66%	19	67%
	2	20	72%	28	73%		15	4	61%	20	67%
S_2	16	21	72%	27	72%	H_2	21	24	74%	21	68%
S_3	30	22	74%	30	76%	H_1	9	29	78%	22	68%
S_4	18	23	74%	29	75%		31	9	64%	23	69%
H_2	21	24	74%	21	68%		25	7	64%	24	70%
	20	25	74%	1	61%		22	14	69%	25	71%
H_4	17	26	75%	15	66%		5	27	76%	26	71%
	5	27	76%	26	71%	S_2	16	21	72%	27	72%
S_1	1	28	76%	31	83%		2	20	72%	28	73%
H_1	9	29	78%	22	68%	S_4	18	23	74%	29	75%
H_3	27	30	79%	18	66%	S_3	30	22	74%	30	76%
	23	31	83%	16	66%	S_1	1	28	76%	31	83%

Table 6.11. Top 15 agreed upon Gaelic folktunes ordered by adjective-pairs (on left, labeled *Adject*) and picture pairs (on right, labeled *Pict*). The left-most column in both sections indicates the tune's affective rating on the *Happy/Sad* scale from Experiment Ten.

6.6.5 Discussion

The results support the hypothesis that subjects demonstrate concordance between two complex stimuli sets. It is only possible to speculate as to the attributes present in the pictures responsible for subject choices for any one tune, much less those attributes in common in the correlations between picture sets or present in the principal components.

Different sets of pictures, such as pair *1/2* and *3/4* bear some evident differences. In both pairs, the first picture is of a snowscene while the second has no snow present; the first picture is mostly composed of blue hues while the second has a wider spectrum. The two picture sets are also well correlated (ρ = 0.70; see Table 6.8). Attempting this sort of post hoc analysis of the picture set characteristics is unnecessary at this point. Subjects in this experiment could be matching any of a number of attributes, but the pattern of results demonstrate that regular associations *can* be made with visual information as rich or richer in content than music.

The top eight tunes from affect rating in Experiments Ten and Eleven figure in the top 50% of rated tunes, whether the tunes are rated by adjective pairs or picture pairs. The qualities in the tune that make it less ambiguous on an affective scale also makes them less ambiguous when measured with adjectives and pictures. As discussed in the last experiment, Osgood et al (1957) hypothesised three main factors employed in the semantic differential task: evaluation, potency, and activity. The higher average concordance ratings of tunes with high average affective ratings this experiment could be predicted if evaluation is a principal component in the picture task, and evaluation is also correlated with the *Happy/Sad* scale.

The four tunes that were rated as the saddest in Experiment Ten (Tunes 1, 16, 18, & 30) were all among the five highest average concordance ratings in Experiment Thirteen. The highest happy-rated tunes do not appear as prominently in the highest average concordance ratings in Experiment Thirteen, although they have high average concordance ratings in Experiment Twelve. Some qualities in the sad tunes are more easily translated onto the SoCs formed in the picture task than the happy tunes. The qualities of the presented pictures lend themselves to more regular matching with the happy tune qualities.

The increased number of factors in Experiment Thirteen revealed in the principal components analysis indicates a more complex utilisation of SoCs in matching across sensory modes. The nature of using adjective pairs in the 2AFC matching task may inherently limit the number of underlying factors available to form cross-modal matches. Increasing the complexity of the matching stimuli while keeping the music at the same complexity means that subjects can use more characteristics of the stimuli to match with the musical qualities.

6.7 Conclusions

Gaelic folktunes, presented without expressive elements or human performance characteristics were shown to be effective in communicating common content to many people. The tunes presumably varied in efficaciousness of communication. The different methods of examining communicated content reveal that those that are less ambiguous in one scale are often among the less ambiguous along another scale.

Experiment Ten demonstrated that the affective scale of *Happy/Sad*, as a continuous scale, could be used to rate the tunes, and that a subset of tunes receive extreme ratings on both sides of the scale. From these results, a subset of four tunes rated highly as *Happy* and another subset of four tunes rated highly as *Sad* were extracted for use in Experiment Eleven. These tune subsets were also used in qualitative analyses for Experiments Twelve and Thirteen to determine if tunes with high concordance ratings also have high affective ratings.

Experiment Eleven examined segments of tunes composed of varying number of beats at the beginning and end of each tune. This was performed to determine if communicated affect was a function of beat length, and to explore what types of musical structures might communicate affect. Separate rating scales for *Happy* and *Sad* were also used, to determine if the scales could be used independently by subjects. It was observed that a *happiness* quality was communicated quickly for those tunes in the *Happy* subset, while *sadness* took longer to establish. *Happiness* was effectively communicated at four or more beats, exhibiting a step function of ratings between two and four beats. This step function was also present in the *Happy* ratings of the *Sad* subset of tunes. It was also found that the *Happy* and *Sad* could be used independently by subjects, as they demonstrated significant ratings on opposite poles of the scales for the same fragment of music. Several musical constructs were suggested as qualities that increase *Happy* ratings: syncopated rhythms, large intervallic leaps, and defined major tonality.

Experiment Twelve used a modified 2AFC semantic differential task to examine the full set of Gaelic folktunes. Subjects rated pieces by selecting an adjective from a pair that represented ends of a continuum. Concordance ratings produced from the percentage subjects agreeing on the same adjective revealed that the music communicated qualities that correlated well with some of these scales for the subject sample. Many of these adjective pairs were those used to describe personality characteristics. The tunes which produced the highest *Happy* and *Sad* ratings in Experiment Ten were among those with the highest average concordance ratings. Thus, those tunes that were least affectively ambiguous were also least ambiguous on the measured categories.

Experiment Thirteen pushed the 2AFC matching paradigm to an extreme by presenting complex pictorial information to match with the tunes. Results indicated that the communicated content of the tunes could be regularly matched with some of the picture pairs. The tunes that best communicated *happiness* in Experiment Ten also had qualities that matched well with the pictorial stimuli. The exact nature of the underlying qualities is not important: only that subjects do extract common information from two sets of complex stimuli and match them in a similar manner. The increased number of principal components when compared to Experiment Twelve indicates that the use complex stimuli may examine more of the qualities present in the music than the adjective task.

The results of the four experiments together indicate that a variety of inquiry methods can be used to obtain a measure of information communicated by music, even in its most mechanical form. The scales may be greatly varied and still be used to examine the music; it seems to be important only that the choices presented in the 2AFC task represent end of an evaluative continuum. Subjects can then align the qualities from the two stimuli together. If enough of the sample extracts common qualities and aligns them in the same way, the result will be significant matching.

7 Synaesthesia Observations with Six Different Synaesthetes

7.1 Overview

This chapter examines some of the synaesthetic correspondences of six synaesthetes available for interviews and experiments. The synaesthetes were five females and one male, each with different sets of cross-modal correspondences. The synaesthesiae included coloured hearing, coloured and geometric pain, coloured words, coloured alphabets, number forms, and coloured numbers. From observations gathered from the subjects in the course of experimentation and interviewing, four recurring features of synaesthesia were identified. 1) Each synaesthesia was fixed and constant in its correspondences; 2) The synaesthetic imagery was involuntarily and consistently evoked when the subject was presented with the paired stimulus; 3) The correspondences were categorical in nature, and; 4) The more potent visual synaesthesiae were three-dimensional in appearance, with the cardinal axes tied to different properties of perceptual experience.

7.2 Introduction

This chapter, along with the following two chapters, deals with synaesthesia and its relationship to cross-modal perception. Six synaesthetes who made themselves available for research appear throughout these chapters. These subjects were available for different periods of testing, which made it possible to record many observations and create specialised experiments to explore their specific forms of synaesthesia. Chapter Seven is composed of observational information gathered in conversations with the synaesthetes themselves, as well as items noted in the course of experimentation. Chapter Eight is composed of case studies using customised experiments devised for some of these synaesthetes. Chapter Nine presents two more formalised experiments, carried out on three coloured-number synaesthetes and controls, to examine how the well-known Stroop effect might be used to examine synaesthetic associations.

Many of the following observations rely upon each synaesthete's self-reporting of past events and introspection into their synaesthetic experiences. There are problems inherent in any introspective technique, but no other option exists for assessing imagery experienced by any individual. The synaesthesiae of the subjects in this chapter have some common features, which will be discussed in detail:

1. Each subject's synaesthesia is fixed in its sensory associations. Subjects repeatedly report the same stimulus/imagery associations whenever asked;
2. The associated imagery are evoked involuntarily, and this imagery impinges on the attention of the synaesthete;
3. The synaesthesiae are categorical in nature;
4. The imagery are often three-dimensional, with cardinal axes often tied to different spatial qualities.

7.3 Synaesthetic Subjects

A small group of synaesthetes was available for study during the process of research for this thesis. The subjects were five female and one male synaesthetes, with potency of synaesthetic perception. The subjects were:

1. DS: a 42 year old female, with coloured hearing and synaesthetic forms;

2. AL: a 20 year old male, with colours associated with the alphabet, numbers, days of the week, months, and other categorical information;

3. MG: a 41 year old female, with coloured alphabet and coloured numbers;

4. AJ: a 41 year old female, with coloured hearing, coloured alphabet, and coloured numbers;

5. CS: a 45 year old female with coloured pain, coloured alphabet, and coloured numbers;

6. LH: a 33-year-old female, with coloured numbers only.

Table 7.1 is a matrix showing the different synaesthetes described above, and the different types of cross-modal associations each of them experience.

Synaesthetes											
Subject	Colour							Synaesthetic Forms			
	Hearing	Letters	Numbers	Days	Music	Months	Pain	Seasons	Dream Shapes	Pain
DS	•	•	•		•				•	•
AL		•	•	•		•		•		
MG		•	•							
AJ	•	•	•		•					
CS		•	•				•			•
LH			•							

Table 7.1. Chart showing the different types of cross-modal associations experienced by the synaesthetes in Chapter 7.

With the exception of DS, all of the synaesthetes were self-reported synaesthetes without formal diagnoses. Cytowic (1989) extensively documents DS's synaesthesia and neuropsychological deficits. DS was flown from the United States to participate in experiments in the course of filming the BBC documentary *Orange Sherbet Kisses* for the *Horizon* programme (1995). All of the remaining subjects made initial contact after either hearing about on-going research through a third party, or through their viewing of the "Horizon" programme. A basic questionnaire was assembled before the transmission of the programme in order to deal with the elimination of potential false self-diagnosed synaesthetes. The questionnaire was short and open-ended in nature, to allow the individual to expound on their particular experiences.

The questionnaire consisted of the following questions:

1. How did you learn that synaesthesia research was being done at Stirling University?

2. What types of associations between senses do you experience?

3. How long have you known that you experience these associations?

4. Do any of your relatives also experience associations between their senses?

5. Have your associations changed over time, or have they always been the same?

6. Have you had any other unusual experiences?

Some responses to these questions were sufficient to rule the person out as a potential subject for testing. As a rule, potential subjects were not considered if they used the screening process as an opportunity to discuss additional experiences that were religious or parapsychological (*e.g.*, extra-sensory perception, telekinesis, prognostication, etc.).

These synaesthetic subjects all experienced colour as a part of their imagery experiences. Two subjects reported experiencing visual shapes in conjunction with specifically musical stimuli (DS & AJ). Three subjects reported coloured-numeric associations (LH, AL, & AJ). Two of the three coloured-numeric synaesthetes also reported coloured-lexical associations (AL & AJ).

7.4 Fixedness of synaesthetic associations

A fixedness between a stimulus and associated imagery has been one of the key criteria in most assessments of synaesthesia. Cytowic (1989) made it one of his five diagnostic criteria, and Harrison & Baron-Cohen used it as a key element of their *Test of Genuineness* for evaluation of synaesthetes (see Chapter Two). Luria's account of the mnemonist S details synaesthetic images that remained constant throughout S's life (Luria, 1968). The synaesthetes that participated in research for this thesis all reported having consistent imagery as far back as they could recall.

7.4.1 The nature of DS's synaesthesia

DS's synaesthesia was primarily visual-hearing, with accompanying synaesthetic forms. Her synaesthetic percepts were the most vivid of all subjects interviewed. Her synaesthetic experiences were strongly integrated into her personality—that is, she was acutely aware that she perceives things differently from most people, and frequently commented on what she was experiencing. Whereas the other synaesthetes might have gone years without knowing that they perceived the world in a different fashion, the idiosyncratic nature of DS's sensations and associations were well known to her and those around her during her childhood. She reported that her synaesthetic sensations caused her great interference in a variety of cognitive processes.

DS visualises images on a "screen"; an imaginary projection area about 18" wide and 12" tall located about 12" in front of her face. She constantly experiences images on this screen, but reports that these images do not obscure any vision, although they regularly compete for attention with the visual scene before her. DS experiences images evoked by most musical stimuli, but experienced the most vivid imagery with those sounds that she found most appealing. Sounds were primarily described in terms of their specific textures and movement on her "screen" rather than hues of the objects. Most images on her screen move on the fronto-parallel plane; those that she enjoys most also move *in* and *out* from her viewpoint; along a third dimension.

DS reports experiencing these favoured percepts frequently when listening to jazz music, especially those with specific instrumental timbres. She responds strongly to the timbres of a recorded vibraphone and synthesizer patches with bell-like tones (*e.g.*, the Pat Metheny composition "The First Circle"). The visual textures evoked by both of these timbres are described as "golden balls", and their movement is described as "falling down the screen." The actual instruments that produce these timbres are physically metal themselves. It is possible that an association between the metallic objects producing certain timbres became the basis for DS's synaesthetic imagery.

DS reports having the same visual images accompanying different types of sound throughout her entire life. When listening to different passages of vibraphone, recorded or digitally produced with a synthesizer, she always responds with descriptions of metallic and spherical images. Ultimately, the synaesthete's accounts for the constant nature of their imagery must suffice when there is no opportunity to test repeatedly over several months or years.

7.4.2 The nature of AL's synaesthesia

AL reports having coloured-numbers and coloured-letter synaesthesia for as long as he can recall. He describes his imagery as "coloured bar-codes"—numbers in colour that he visualises in the space in front of him. He describes these images as appearing without a background. He says only that they are located before him "in space." The colours that are associated with the different symbols have been consistent throughout his life. In all of the testing, his descriptions of the images were always identical, and he could immediately visualise them on command for description.

AL's coloured-alphabet was recorded using two methods. He first coloured the letters on paper with coloured pencils; then we worked together to get as close of a match as possible using the NeXT computer's colour interface and VDU. The coloured-alphabet worked out with the NeXT is displayed below as Figure 7.1.

Figure 7.1. AL's coloured-alphabet, displayed in approximated matches using computer generated colours.

AL perceives these colours regardless of the presentation method of letters or numbers. He reports experiencing coloured-sensations when hearing a number read aloud, looking at it on paper, or simply thinking of a particular number. The strongest sensations are experienced when viewing printed characters. The coloured-sensations are experienced regardless of font or point-size, and are not reported to degrade with exposure. AL reports that he experiences levels of grey for all punctuation, also described as varying in terms of brightness. For example, commas and quotations are coloured medium-grey, and full-stops are solid black.

One explanation for the increased potency of synaesthesia with printed material may lie in the fact that this would be the most common mode of presentation throughout AL's life. Colouring of letters and numbers is almost universal in children's toys and primers. In the course of development, AL may have persisted in using the SoC of different colours to discriminate between and categorise different characters, while other children go on to abandon this relationship. The residual effect would be an evoked sensation of colour when viewing or thinking of a character.

Another example of the fixedness of AL's synaesthetic associations occurs when he discusses his imagery with other synaesthetes. As was noted in Chapter Two, the imagery experienced by synaesthetes is idiosyncratic, so that any two synaesthetes with the same type of synaesthesia (e.g., coloured-letters or number-forms) will describe very different imagery for those experiences. When shown another synaesthete's coloured alphabets, AL firmly states that they are "simply wrong", and cannot understand how such associations make sense to another.

When he met another synaesthete who experienced synaesthetic forms for days of the week, he said that he could objectively see the logic in her representations, but it was also "incorrect" for him. This synaesthete's rejection of another synaesthete's correspondences is ascribed to the differences that arise when comparing another set of fixed imagery to his fixed but disparate imagery.

7.4.3 The nature of AJ's synaesthesia

Subject AJ also has coloured alphabet and numbers, as well as shapes in response to music. Like AL, her alphabet and numbers have had the same colours throughout her lifetime. She does not report having any "screen" of display like DS and AL, but describes the coloured letters and words as being "just there," in front of her.

AJ experiences different images when letters are formed into words, or when single digits are combined together to form multi-digit numbers With many of the words that she describes, the word has an outline of another colour around it. Some synaesthete's report that numbers or letters glow with an aura (e.g., AL's number 1). For AJ, the "outline colour" is a solid colour that surrounds the word, letter, or number. Figure 7.2 shows AJ's representation and colouring for the digits *1* and *8* separately, and combined together into the number *18*.

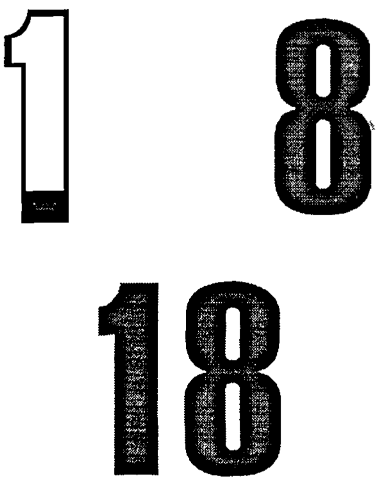


Figure 7.2. Subject AJ's representation of the separate numbers *1* and *8*, and the representation of the two digit number *18*.

AJ experiences coloured words where the colour changes from left to right across the entire word. Her associations appear to be based on the first letter of a word rather than the phoneme. When presented with the words "fish" and "photo" she reports different coloured words. In the case of "fish" and "fire", both of the words share an overall colour. The two words do not evoke exactly the same imagery, as the other letters cause the word to change in colour along its length. The overall word colour taken from the first letter of the word contrasts to the way that her multi-digit numbers are coloured, as they are primarily based on the final digit.

AJ, in common with most coloured-number and coloured-alphabet synaesthetes, uses her correspondences as a way of categorising and remembering things (Cytowic, 1989; Luria, 1968). In terms of categorisation, AJ needs to use one of her colour correspondences as the primary category. It is asserted that for AJ, the right-most digit of a number and the left-most letter of a word form primary categories. The overall synaesthetic percept is coloured by this primary category.

A strong piece of evidence for the fixed nature of AJ's correspondences is the repeated production of her coloured alphabet and days of the week. When in Stirling, she drew representations of these using the coloured pencils available. When contacted later, she remarked that she was unhappy with the quality of the result, and it was suggested that she produce the set again using paints. She said that the new paintings had more precise representations of what she visualised. These pictures showed the same colours for each element, as well as the same changes in colour along the surface of the words.

7.5 Involuntary nature of synaesthetic associations

Each synaesthete interviewed described their synaesthetic percepts as appearing involuntarily whenever presented with the source stimulus. This facet of synaesthetic perception has been one of the few qualities that researchers concur needs to be present for cross-modal imagery to be termed as synaesthesia. With the synaesthetes researched for this thesis, the imagery was never reported to disappear or degrade with repeated presentations of the source stimulus. AL reported that the colour for a number or letter was equally strong regardless of number of presentations. DS always experienced her images of falling metallic balls regardless of how long a recording was played for. This suggests that the synaesthetic correspondences are deep, fundamentally connected to their source.

7.5.1 Attention demanding nature of DS's "Dream-shapes"

Amongst her synaesthetic percepts, DS experiences "spontaneous synaesthetic forms" during times of stress. She describes these as *dream shapes*, that she visualises when falling asleep. She describes the imagery in terms of pleasantness, the most pleasant being "bolts of fabric unwinding" and the most unpleasant as "...metal ball-shapes descending in a group. They disintegrate as they fall, and when they reach the bottom, you know you're dead." She experiences the later only a few times during the year, and reports that she *must* attend to the images. She reports that the dream images, like moving imagery on her "screen", cannot be ignored or disregarded.

DS also reports that images during normal activity demand her attention, and can interfere with the task that she is concentrating on.

7.5.2 DS's performance in cross-modal experiments

DS took part in the six experiments from Chapters Three and Four to examine how a coloured-hearing synaesthete performed in a cross-modal task. These experiments and the results are discussed in detail in Chapter Eight. The behaviour exhibited by DS, and the information provided by her in the debriefing, shows the attention-demanding nature of her synaesthetic percepts. In the course of the experiment, DS would look away from the screen at the onset of every audio stimulus, and did not look back at the VDU until the note(s) had finished sounding. During debriefing, she reported that her attention was drawn to her "screen", and that she had to look at her image "until she understood" what she was visualising. At that point, she could look back to the screen and compare the presenting image with her percept. DS stated that it was *impossible* for her to look at the screen in the presence of her imagery.

7.5.3 Attention-demanding qualities of AL's "Bar-codes"

AL also reports that synaesthetic imagery is *always* generated by numbers and letters, and he must attend to the imagery before he can attend to the content of the evoking stimulus. In the course of experimentation, AL was verbally and visually presented with letters and numbers in memory tasks. Whenever presented with a stimulus, his eyes would move automatically to his upper-right visual quadrant, then back to the experimenter or the stimulus. AL reported that he "had to look at the numbers" or other imagery when it appeared. For both AL and DS, the attention-demanding quality of their imagery actually drew their gaze away from its previous focus.

7.6 Categorical nature of synaesthetic associations

The synaesthetes researched for this thesis were all selective synaesthetes with the possible exception of DS. As discussed in Chapter Two, it is hypothesised synaesthesia is generated from the act of using information from other modalities as SoCs, and persisting in the use of these SoCs. Finding categorical synaesthetic associations in the mature individual, then, is expected. Some of these categories are obvious, such as colour associations with numbers and letters. Others are more abstract to anyone other than the synaesthete herself. For the individual, however, categories may be internally consistent. The experimenter must rely upon the introspection of the synaesthete herself to help infer the categories of stimuli that evoke associated imagery.

7.6.1 Colour-Number Correspondences

Coloured numbers are a relatively straightforward set of categorical stimuli. Of the six synaesthetes available for research, five had coloured numbers (all with the exception of DS). Each had their own idiosyncratic colours, but showed a commonality in the detail they gave to the digit "1". The synaesthetes all took great care in describing "1" as fully as they could, and frequently experienced difficulty in adequately explaining what they were visualising. AJ describes "1" as a white digit with a blue outline of colour surrounding it (see Figure 7.2). For Subject LH it was described and drawn as white with an outline. Subject AL reports that he has difficulty accurately describing "1", but the closest description he could achieve was "...very light, like white, with a blue glowing aura surrounding it." The relatively detailed image for "1" compared to the remaining digits may be due to the fundamental nature of the digit. The recurring use of the colour white in the images suggest a fundamental correspondence between the most basic of numbers and the primary quality of white. One paradigm for the order in which colours are learned was proposed by Berlin And Kay (1969), with white and black learned first, followed by red, then the remaining colours.

Other descriptions of synaesthetic imagery for "1" are given in existing literature. In Cytowic's reports (1989), Subject JM describes "1" as white with a black outline (p. 38), Subject GG describes it as simply "black" (p. 39), Subject SdeM sees "grey-black" (p. 210) and Subject MT sees "white" (p. 228). There is a primary quality of colour in each of these associations. Even with the disparity in associations of every other alphabetic or numeric character, it is notable that each of these reports involve "white" or "black" in their descriptions. Their associations for that particular character were extremely specific in nature.

7.6.1.1 *Basis for Colour-digit Associations*

Tracing synaesthetic correspondences back to childhood associations has met with the criticism of being too simplistic (Cytowic, 1989). Although certainly simplistic in its basis, it may still be valid. Synaesthetic correspondences must have *some* origin if they can be categorical in nature. An early association with a particular toy or book may form an initial association that can develop into a permanent synaesthetic correspondence.

These types of associations have been hypothesised, but no mention of evidence for an early association evolving into synaesthetic perception is present in the synaesthesia literature, except where such evidence is taken from the synaesthete's own memory. However, two recent experiences have provided some evidence for traceable developmental associations.

Subject, LH, who experiences coloured number synaesthesia, is the mother of identical twin boys, RH and TH, aged four. LH reports that RH is especially attentive to the colours of objects, freely sorting objects primarily by colour when playing. During a session with the University Creche instructor when naming digits presented on a computer screen, RH was asked what number he saw on the screen. Although the number 8 was being displayed on the screen, RH responded without hesitation, "Orange!". When asked again, he repeated his response.

In the course of discussing this incident, it was revealed that RH and his brother both spent much time playing with a coloured number puzzle. RH's parents asked him for his colour-digit correspondences and recorded them. The puzzle was compared against RH's responses and all colours matched exactly. It is still in question, of course, as to whether RH will develop into an adult synaesthete.

In addition to the colour-number confusion, RH's parents report that he gives a large degree of attention to the colours of all objects that he plays with. It may be that this extra attention given to the intrinsic quality of "colour" is one of the underlying factors for the development of synaesthesia. The same learning paradigm that explains coloured numbers can apply to coloured alphabets.

The coloured alphabets of Subjects AL and AJ have already been outlined in Sections 7.4.2 and 7.4.3, respectively. Subjects CS and MG both experienced coloured alphabets as well, and provided coloured drawings of their images for each letter. In accordance with existing literature (Calkins, 1895; Whipple, 1900; Marks, 1975; Cytowic, 1989), the colours were idiosyncratic in their correspondences.

One common feature of coloured alphabets noted by Baron-Cohen et al (1993) is the number of synaesthetes that pair the letter "O" with white. They point out that a majority (73%) of recorded alphabets (Galton, 1883; Jordan, 1917) in conjunction with their subject pool show this correspondence ($N = 33, p < .001$). Of the coloured alphabet synaesthetes, only AL has this correspondence; AJ sees "O" as orange outlined with red, and CS sees it as yellow.

7.6.2 *Other categorical synaesthesiae*

Other categorical synaesthesiae were observed: synaesthetes AL and AJ had correspondences between days of the week, and subject AL experienced seasons and time schedules as synaesthetic forms. For these correspondences, as with letters and numbers, the categorical nature of the synaesthesia is readily apparent.

7.6.2.1 *AJ's coloured days of the week*

AJ experiences coloured images for all words, so it is necessary to determine if her colours for days of the week are governed by the component letters for the day's name, or by the abstract concept of the day itself (*e.g.*, some type of "Monday-ness"). When asked about this, AJ replied that it was the days themselves that were matched with the colour patterns. Her images for days were also surrounded with aurae, and do not change in colour along their length, unlike the imagery for regular words. The colours associated with the day names were also different from the first letter of the word itself, also distinguishing them from the regular word images. These associations demonstrate how identical items (*i.e.*, written words) can produce different types of imagery when the different items correspond to multiple forms of categorical information.

7.6.2.2 *AL's Categorical Synaesthetic Forms*

AL has different forms for the representation of serial elements (numbers) and temporal elements (days of the week, seasons, etc.). Synaesthetic literature refers to these types of associations as *synaesthetic forms* (see Chapter Two). The dimensional nature of these images is discussed in detail in Section 7.6.3.3. AL experiences imagery for days of the week, seasons, and time schedules.

7.6.2.3 Importance of brightness in AL's perceptions of material without existing synaesthetic correspondences

Whenever called upon to describe his synaesthetic percepts, AL would speak in terms of *brightness* before turning to other adjectives, especially when he found adequate descriptions difficult. In an exploratory fashion, AL was presented with non-roman alphabet characters to examine any synaesthetic-type responses to the novel characters. These included Greek, Cyrillic, and Hebrew characters, mathematical symbols, and simple geometric figures. In all of these cases, AL's reported that his percepts could only be described as varying degrees of brightness. Those items which evoked synaesthetic-brightness responses in AL were described confidently in terms of their brightness.

Using the synaesthetic development paradigm, AL was still employing the SoC of varying brightness to categorise different characters. Additionally, the associated brightness sensations for novel symbols may initially be similar in nature to those symbols with colour-associations that they most closely resemble in shape. For example, the colour that AL reported experiencing for the Hebrew character aleph (א) is quite similar to his colour-assignment for the letter "X", and his experience for capital sigma (Σ) is similar to the number 3. It seems likely that if he engaged in learning a foreign language with a new character set he would develop a full set of synaesthetic associations for the characters, although these sensations might be less potent in quality.

7.6.3 The three-dimensional qualities in synaesthetic correspondences

Synaesthetic correspondences often have a three-dimensional quality. Some have been touched upon already—for example, the spherical shapes of DS's images for vibraphone. In addition to imagery being three-dimensional, the axes of those dimensions may be tied to different qualities of the stimuli that evoke the synaesthetic imagery.

Not all of the synaesthetes experienced three-dimensional imagery. The potency of recorded synaesthetic experiences varied between descriptions of highly detailed three-dimensional objects with textured surfaces to descriptions of a simple "sense" of colour. The detail of imagery is an indication of how strong the synaesthesia is for that person. The examples given below are of three-dimensional imagery from synaesthetes with strong correspondences. More examples are given for AL's associations, as he was available for testing over a one-year period.

7.6.3.1 The three-dimensional qualities of DS's music imagery

DS's description of the imagery that she experiences when listening to music is very detailed. Most often, she describes movement across her "screen" in two dimension, moving within the plane of her screen. She reports that trumpets occur high on her screen, possibly tied to frequency height and clarity of timbre. The movement of "dropping golden balls" may be related to the short duration of the tones, and the decay of those notes. Although it can be observed that the dimensions of movement and placement on the screen are tied to different qualities in the music, it is only speculation as to what exact qualities are being represented.

One dimension of the music that is represented in her imagery is her *affective* response to the music. DS herself reports that the music that she enjoys the most moves in a third-dimension moving towards and away from her. Exactly how the images move along this dimension are unknown, and ineffable for DS.

DS's "dream shapes", outlined in Section 7.5.1, also have a high affective component, and move in three dimensions. The experience of seeing these images is either highly enjoyed or literally dreaded by DS. It is again likely that the movement of these shapes is related to changes in DS's internal states, but it is difficult to assert any more than that.

7.6.3.2 The three-dimensional qualities of CS's pain forms

Synaesthete CS reports having strong imagery that accompanies pain sensations. Many of these she has rendered in paintings, and can describe the sensation that initially produced the imagery. For example, one painting of a white dagger-like image on a coloured texture background is described by her as toothache. CS has turned to sculpture to represent other tactile sensations, as the painting do not adequately capture the three dimensionality of her imagery.

7.6.3.3 The three-dimensional qualities of AL's synaesthetic forms

AL also has three-dimensional qualities in his synaesthetic forms, and can usually identify what the axes represent. One example of this is his synaesthetic form representing a number line. He describes this as a set of numbers images that have thickness, angling up and away to the right from his point-of-view, with his head positioned at "0". When asked to describe negative numbers, he reports these as following the same pattern, but angling up and away to the left of his point-of-view. This bears little similarity to other reported synaesthetic forms (Cytowic, 1989) in its layout.

AL describes his imagery for a schedule of a day's activities as an image similar in layout to a page from a desk planner. He perceives different horizontal slots for activities divided up by blocks of time. Those times with activities are perceived as text with the activities description, and an increased "brightness", to use his description, for that block of time. Those activities that he looks forward are coloured brighter. His memory for events is self-reported to be excellent. Thus, AL has a traceable affective link between his visual SoC of brightness and important events in his memory.

AL uses a similar imagery form for picturing days of the week. He describes the image like a pavement, angling up and away to the right from his point-of-view in the same manner as his number-line. He has a colour for each day of the week, with Saturday and Sunday coloured a very bright yellow, and slightly raised up in position. Figure 7.3 provides an approximation of the imagery he experiences. This figure was produced with AL's guidance.

AL's point-of-view can alternate between the first day of the week or the current day of the week, dictated by context. When planning his week, he sees days that have important events scheduled as coloured brighter than their neighbours, and can see the schedule for the day, visualised as described above, superimposed on the image for the day itself. He says that he can see forward approximately three weeks from the present day. When asked how he pictures days and weeks past, he either pictures them along the same view if they had just passed and he is viewing from the beginning of the week, or sees them angling up and away to the *left* from his point of view (POV). This method of visualisation is almost identical to his imagery for negative numbers.

The dimensions here are tied to the qualities of time and affect. Dates in the future extend away and to the right of AL's POV; those in the past also stretch away but to the left of the POV. Weekend days, which AL reports looking forward to, are raised on the image, suggesting a visual representation for that preference. As mentioned earlier, events that he looks forward to also appear brighter.

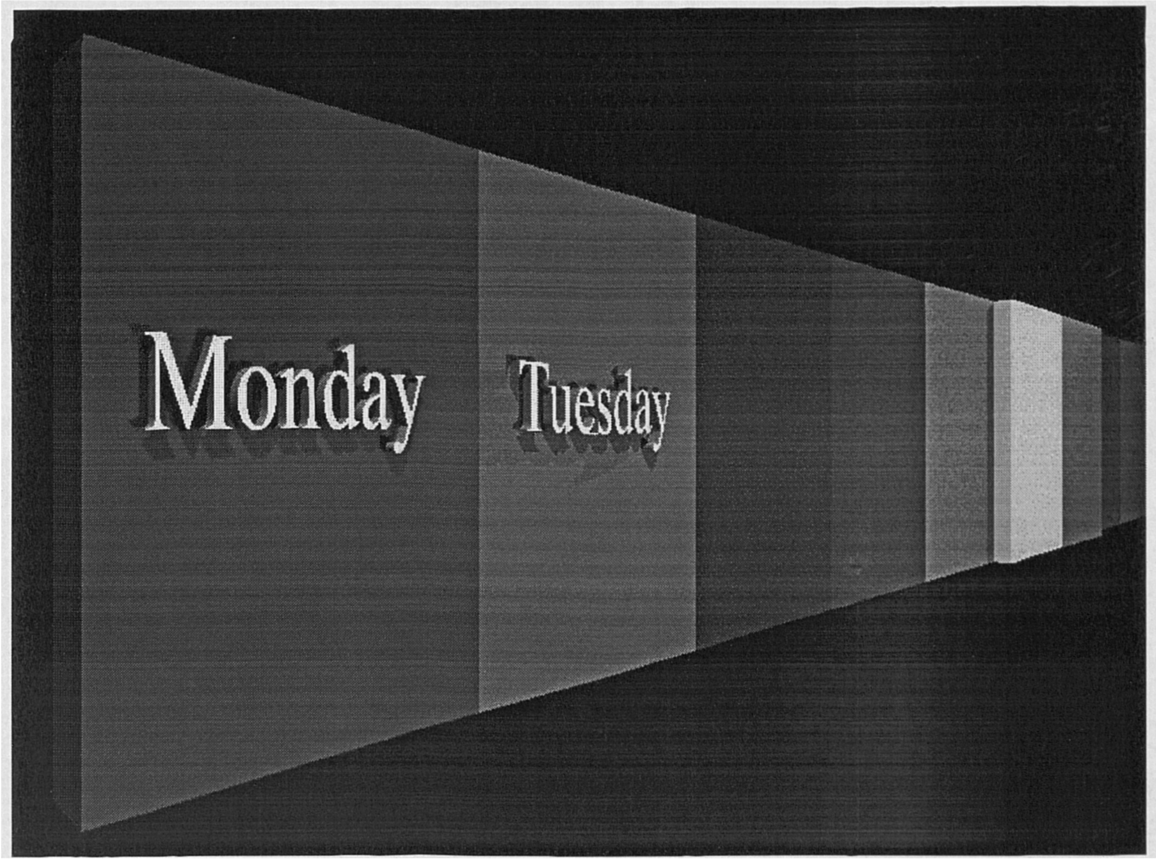


Figure 7.3. Approximation of AL's synaesthetic form for one week in the future. AL is positioned at Monday (Red rectangle). The weekend days are slightly raised and in shades of yellow.

Finally, AL uses yet another schema for seasons. He pictures the year as "...I am a small man standing on an oblong clock," with himself positioned on some part of the clock. This oblong shape is also reported as three-dimensional, stretching away from him. His POV only allows him to view part of the entire image at a time. Things that he reports looking forward to, such as Christmas, Summer, and term breaks are not specifically coloured, but instead described again as being "brighter" than other areas and slightly raised. As with visualisations of weeks, he could relocate himself along the image when imagining himself in the past or future. The seasonal representation has dimensional representations of time and affect similar to the imagery for days of the week. Here the movement across the perimeter of the shape is temporal, and the vertical displacement again is an indication of affect.

The internal consistency of AL's synaesthetic forms is more apparent than those in many other reports. Fundamentally, the employment of these forms by each synaesthete is the same: visuo-spatial representations of conceptual categories are used as schemata for cognition. An objective evaluation of a particular form's effectiveness is unimportant: it is only necessary for that form to represent information in a manner understood by a synaesthete. Number forms may be any shape, such as zigzags and irregular curves (Galton, 1907; Cytowic, 1989). It is possible that some representations will lend themselves more readily to different tasks, such as AL's superior short-term memory abilities (see Chapter Eight).

AL provides a much more orderly set of associations in his synaesthetic forms than other synaesthetic subjects in this thesis and in many existing reports. This orderliness is the type that might be devised by a non-synaesthete attempting to model synaesthesia rather than the more complex and seemingly random forms commonly reported. The important attribute for a synaesthete's perceptions is that the associations are regular for her; the continued use of the forms reinforces them and structures the organisation of the synaesthete's memory. This constant organisation can account in part for the reported increase in the synaesthete's memory capacities. Also, examination of synaesthetic forms allows an outside observer to view a direct representation of some of the synaesthete's internal organisation and cognitive processes. Understanding of which correspondences best lend themselves to increased memory capabilities may aid in the development of early teaching programmes and aids.

7.6.3.4 Spatial characteristics of AL's "bar-codes"

AL reports that numbers are visualised closer together than letters. It is possible that the presentation of letters as separate entities affects the spacing. He reports that a visualisation of a word has the letters closer together than when the same letters are presented separately. Additionally, spacing changes for single digits and those same digits assembled as a multi-digit number.

This change in spacing can be controlled by AL. One demonstration of this are the images when AL was asked to picture the separate digits "1", "4", "5", and "9", then to picture the single number "1,459". AL reports that the same bar-code pattern appears for each; however, the latter image was closer spaced. Thus, the dimension inter-element spacing represents whether elements are thought of as individual elements or as part of a larger item. Further evidence for the context effect on the imagery is the way that a set of six or seven individual digits can become more spatially compact when AL regards them as a telephone number.

7.7 Other observations of synaesthetes

A side-note regarding the affective components of synaesthesia regards the responses of several subjects to viewing physical representations of their synaesthesia. Subjects AL, MG, and AJ all spontaneously commented how pleased they were to see the representations of their images on the computer VDU and drawn on paper. Both AL and AJ made very similar comments, in that the representations made them feel comforted. When seeing the coloured alphabet on the VDU, AL remarked that it felt "...like something that I have created." One correspondent wrote in that seeing the coloured concluding titles of the Horizon programme made her feel safe. DS remarked that she greatly enjoyed seeing representations of music that approximated her synaesthesia.

From the different subjects interviewed, part of these feelings might be ascribed to an affirmation that external, physical manifestations of the synaesthete's idiosyncratic correspondences can exist outside of their internal representations. Synaesthetes commonly report feeling isolated in their perceptions, and frequently report ridicule or disbelief when discussing synaesthesia with others. The synaesthetic correspondences are fully integrated into their lives, and several report that years passed before they were even aware that others did not perceive the world in the same fashion as they. This begs the question of how many individuals have synaesthesia and are as yet unaware of it, as they have never considered that others may not share in their modes of perception.

7.8 Conclusions

The synaesthetic sample available for research for this thesis was small, but many qualities of their different synaesthesiae concur with existing accounts of other synaesthetes. The idiosyncratic nature of these experiences make standardised synaesthetic tests almost impossible. This necessitates the case-study approach for examining synaesthetic percepts.

Each of the four main features of synaesthesia outlined here can be ascribed to the hypothesis that synaesthesia develops from an application of existing SoC when encountering new types of information. The fixedness in associations arises from the consistent application of the same SoC, which forms the synaesthetic association itself. The involuntary nature of the imagery evocation arises from the manner in which the element (*e.g.*, number or letter) and its associated correspondence together form a schema, so that the element cannot be imagined in isolation. The categorical nature of the associations arises due to the manner in which a SoC was initially applied in order to form schemata. Finally, the spatial qualities of the more potent visual synaesthesiae arises from using the cardinal axes of the space as additional SoCs to represent other qualities, such as temporal or affective qualities.

8 Experimental Case Studies of Synaesthetes

8.1 Overview

This chapter has two sets of experiments conducted with some of the synaesthetes discussed in the previous chapter. It first examines the performance of synaesthetes who have coloured imagery in the cross-modal experiments from Chapters Three and Four. Two of these synaesthetes had coloured hearing; the other had coloured letters. The results demonstrate that some synaesthetes can use the varying grey levels a SoC for matching with musical qualities. However, the fixedness in synaesthetic associations that include colour preclude the synaesthetes from using colour as a SoC in the coloured circle experiments from Chapter Four. The second set of experiments explored the increased STM abilities of synaesthete AL. These experiments revealed that AL's imagery allows him to recall elements serially as well as in arbitrary orderings. Oral and visual presentation of phonologically and synaesthetically confusable letters creates problems in memory encoding. The interference of the letter's phonological and synaesthetic content changes with the mode of presentation.

8.2 Introduction

This chapter contains two sets of experiments. The first set of experiments are replications of the cross-modal experiments found in Chapters Three and Four with coloured-hearing (DS and AJ) and colour-alphabet (MG) synaesthetes as subjects. The second set of experiments are explorations into the increased STM of synaesthete AL, using visually and orally presented strings of letters and numbers.

8.3 Synaesthetic Performance in Cross-Modal Experiments

8.3.1 Introduction

The logical extension of cross-modal experimentation is to compare results with those of individuals who experience concrete and consistent cross-modal perception—that is, with synaesthetes themselves. Synaesthetes provide the opportunity to examine the process of cross-modal associations in stasis. Cross-modal perception involves drawing correspondences between different modalities and concepts as required. The synaesthete shares this ability when dealing with sensations that lie outside of her existing synaesthetic correspondences. The fixedness in existing correspondences should preclude synaesthetes from using the separate qualities as SoCs.

The synaesthetes who took part in the cross-modal matching experiments already used colour and shape in their correspondences. For DS, it was one of a number of qualities making up her florid imagery; for AJ, it was a simpler form of coloured hearing. Subject MG did not have coloured-hearing, but did pair colour with categorical stimuli (letters, numbers, days of the week). Due to the limited time available with AJ, she only took part in the single note matching with grey circles and coloured circles (Experiments One and Four, respectively).

8.3.2 Experiment DS.1: Grey circle matching with simple musical constructs (Experiments One–Three)

8.3.2.1 Results

DS's performance in these three experimental tasks was extremely consistent. In all three of the experiments combined, the brighter of the two circles was selected by DS only 6 times out of all 164 separate trials. No single stimulus received a score greater than one. In the *Tones* experiment, only one tone was matched with a lighter circle. In the *Phrases* experiment, three of the four intervals with the Cs_5 as the target note (TN) had a score of one; all of the G_3 and D_4 chords were matched with the darker circles in all trials. In the *Chords* experiment, only the F-dominant 7th chord received a score of one; all others were zero.

Subjects AJ and MG both had significant regression coefficients for the *Tones* experiment (AJ: Pearson's r = 0.93, $p < .001$; MG: Pearson's r = 0.85, $p < .001$), indicating that they, like the majority of the non-synaesthetic subjects, matched increasing pitch height with increasing visual brightness. Regression of Subject MG's scores and TN in the *Phrases* experiment also reveals a significant coefficient (Pearson's r = 0.87, $p < .001$), showing the same matching direction of pitch height and visual brightness. No regressions on scores in the *Chords* experiment were significant.

8.3.2.2 Discussion

The pattern of DS's responses indicates that she responded consistently by always selecting the darker of the presented circles. One possible explanation for her performance is DS's difficulty in selecting a grey level that adequately approximated what she was visualising. The normal subjects performing the matching tasks had to choose a representation across modalities; that is, they could select characteristics for SoCs and align them in a personally consistent manner. DS's internal representations of music already has visual qualities. For her, these visual qualities are not available for use as a SoC.

Another explanation for her selections comes from her own introspection into her performance during the debriefing. She asked what normal subjects did in the experiment, and was told about the common performance of matching increasing pitch height what increasing visual brightness. As mentioned in Chapter Three, she responded by noting that she felt that the darker circles, closer to black, were "brighter" for her. The lighter grey circles were less "substantial" in their colour, in her opinion. As some subjects did in Chapter Three, DS may have equated increasing contrast with increasing brightness. It is important to remember that DS experienced vivid colour and form associations for each of her images. Since all of her associations are vivid and concrete, DS may have always selected the brighter colour for her.

The performance of the other two synaesthetes is in agreement with the majority of non-synaesthetic subjects for the *Tone* and *Phrase* experiments (AJ did not participate in the *Phrase* experiment).

8.3.3 Experiment DS.2: Colour circle matching with simple musical constructs (Experiments Four–Six)

8.3.3.1 Notes on DS's performance during the experiments

DS performance during the coloured circle experiments was quite distinct from all other subjects. At the onset of each musical stimulus, DS would look up at an angle above the video monitor until the conclusion of the stimulus. At this point, she would blink forcefully and look back to the monitor. She would often complain that she disliked both of the colours, and did not want to select either. Looking away from the screen was much more pronounced for the coloured-circle than the grey level circle conditions. She was re-instructed to select the circle that she felt was most appropriate, and her frustration continued to grow on each successive trial. She said that the *Phrase* experiment actually made her irate and "...angry! I want to put my fist through the monitor." She was surprised to hear during the debriefing that all subjects did not hate the experience of the experiment, and incredulous that some subjects actually enjoyed and were relaxed by the experiment. She reported feeling drained at the conclusion of the experiment.

8.3.3.2 Results

None of the regressions run on any of the subject's scores computed by affective rating, subjective brightness rating, or luminance reveal any significant coefficients. Figure 8.1 shows DS's scores for the *Tones* condition using coloured-circles. In the same manner that the results were computed for Experiment Four, the scores were calculated as a number from 0–5, calculated using different evaluative criteria. The graph shows the scores for each of the tones, ordered along the abscissa by pitch height. The separate bars show the scores by more attractive circle chosen, brighter circle chosen (both computed from DS's affective and brightness ratings for the colours), and brighter circle chosen in terms of luminance.

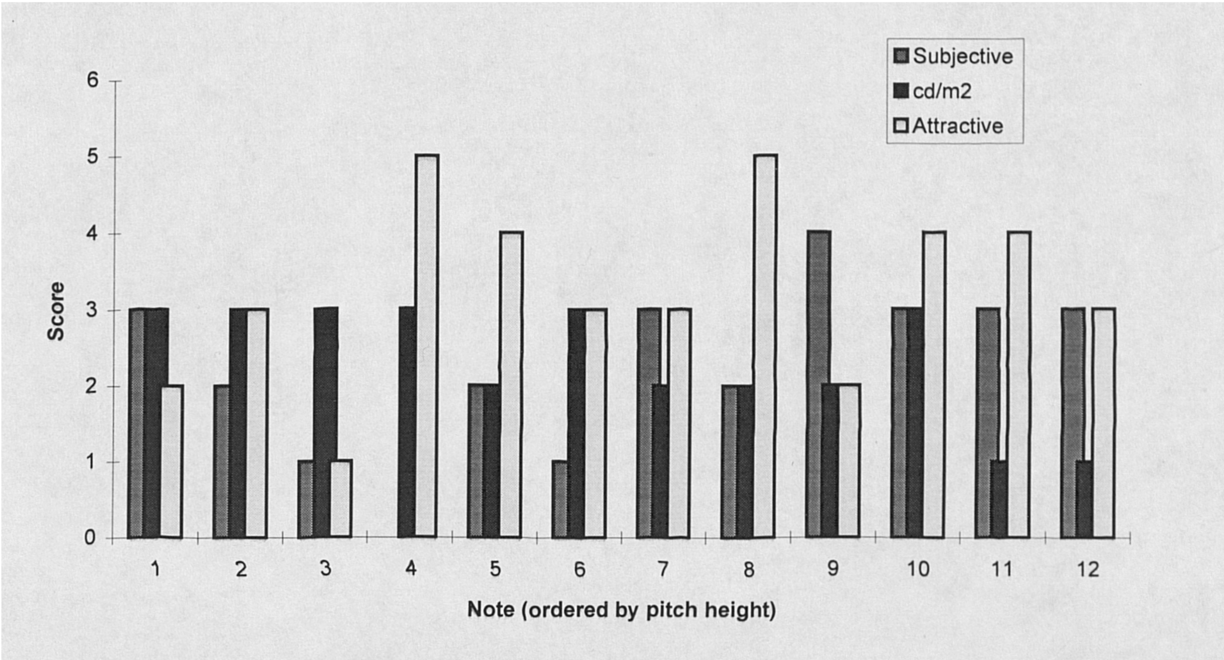


Figure 8.1. DS's scores for cross-modal matching of single tones and coloured-circles computed from subjective brightness, subjective attractiveness, and cd/m^2.

8.3.3.3 Discussion

No clear pattern emerges from DS's responses. It would be more of a surprise if a pattern *were* to be observed with the data analysed in such a manner. Regular matching performance would be observed only if brightness, luminance, and/or attractiveness had been factors used to generate the initial correspondences that developed into DS's synaesthesia.

The lack of significant results cannot prove that the synaesthetes were responding in a manner different to non-synaesthetes, but the comments of DS does suggest that she, at least, was unable to match presented colours with her images. In DS's own words, "The colours are *wrong*; I want to *fix* them." She reports that the shape of the images on the screen were "...very wrong. They were all circles, and none of the sounds were round!" DS's associations are concrete and constant for her; any items that did not closely approximate her visualisations were regarded as equally incorrect. This may also explain the lack of pattern in results for the other two subjects.

Why should there be a pattern of matching for the monochromatic grey circles but not for coloured circles for some synaesthetes? It is possible that grey levels are available for use as a SoC because they are somehow different from colours of varying hues for the synaesthete, although grey levels are technically defined as colours. While the colours are already part of coloured-hearing correspondences, grey levels are not, and thus available for use as a SoC.

8.4 Extended Experiments with AL

Another set of experiments were performed with AL, who experiences a variety of synaesthetic associations discussed in some detail in Chapter Seven. AL came to this researcher's attention during a psychology practical demonstrating STM and digit span. He surprised the class by recalling the entire list of 10 numbers after hearing them once, and casually remarking that the colour pattern of the numbers made them easy to recall.

The historical and modern literature on synaesthesia often comment upon the increased memory abilities of synaesthetes. These abilities may go as far as hypermnesis: the nearly inexhaustible memory of S. (Luria, 1968) was a possible result of his synaesthetic associations. The potential role of synaesthesia in memory is a notable enough feature of many cases that Cytowic (1989) makes this one of his diagnostic criteria for synaesthesia (see Chapter Two).

Psychological testing with AL revealed that he did have memory abilities above the norm, as well as a high IQ. Table 8.1 gives AL's scores on the Weschler Memory Scale (WMS), Weschler Adult Intelligence Scale (WAIS), and Raven's Progressive Matrices.

WMS		
	MQ	118
WAIS		
	Verbal IQ	119
	Performance IQ	128
	Total IQ	125
Raven's Progressive Matrices		
	Part A	10
	Part B	27
	Total	37

Table 8.1. Neuropsychological measurements for synaesthetic subject AL (coloured numbers, letters, days of the week, and seasons).

AL's memory was above the norm; however, his ability in recalling serial lists of elements was remarkable. He could routinely recall lists of 20 elements, and recall these lists several days after presentation. When AL first arrived at the research lab, he recalled the sequence of numbers that were presented in the classroom, learned at a time when he did not know that he would be participating in experiments. Two years later he could still recall the sequence on demand. When asked how he had encoded this sequence so durably after a single exposure, he commented that it was merely the "bar-code" of colour inherent in his associations for that particular sequence that facilitated the memorisation. In the WMS and WAIS digit-span tasks, AL routinely scored maximum in forward and backward recall.

This report, along with exploratory conversations with AL led to the tailored construction and execution of a variety of memory tasks involving numerals and single letters as stimuli. In pilot trials AL demonstrated a digit span of over 10 alphabetical characters and numeric digits. While performing the WMS digit-span task, AL volunteered that he went through an active visualisation when listening to the numbers, and that the task would be easier for him if the numbers were written on paper. Using this suggestion as a guide, the following recall tasks were all presented on A4 white paper, presented in portrait orientation, in Times-Roman 36 point bold font. The letter and digit orders in all experiments were chosen by random selection.

8.4.1 Alphabetical character and digit recall

8.4.1.1 Experiment AL.1: Alphabetical Characters in Columns (2 x 10)

AL was first presented with twenty non-duplicated alphabetical characters arranged in 2 columns and 10 rows (**Col 1**: V A X F Z G N B D S; **Col 2**: S M I P Y J W Q L K). After viewing the list for 60 seconds, AL was asked to reproduce the list in columnar orientation, and quickly did so without error. He was also 100% accurate when recalling the entire list by rows, or recalling individual columns backwards. With a slight delay, he could also recall any row or element in the list by its number; for example, he could supply the second element in the third row after about a 3 second pause. After a nine day interval, AL reproduced the entire list again, in either orientation, with total accuracy. AL's preferred recall orientation was columnar; he recalled in different orientations slightly slower. AL's memory recall appeared to be spatial in nature, as AL would often describe what he was visualising as he was recalling items, and would spontaneously point to positions in the air when recalling in non-columnar orientations.

AL could manipulate the visual images for recall. He could recall the list by giving the first element from column one and last element in column two, the second element from the top for column one and the second from the bottom element in column two, and continuing pair-wise through all the columns. The recalled item performance is supplied in Figure 8.2.

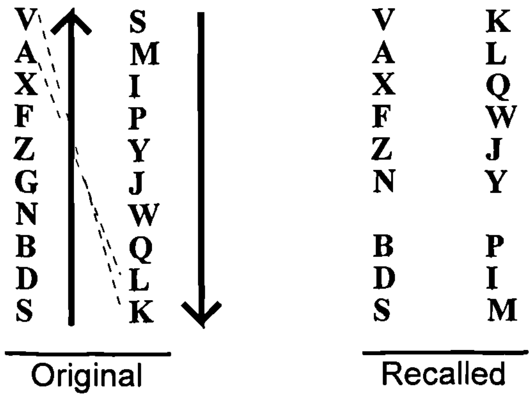


Figure 8.2. Original list and AL recalled list for columnar manipulation. The recall directional order is indicated by the arrows at the right of each original column.

The absence of the letter "G" in the first list provides some insight as to how AL pictures the columns. The elements from the second column are still recalled in serial order, and the final element from the list ("S") is omitted, as AL had run out of elements in the first list. AL also noted that he had made a mistake at some point, as he stated that he still had an element from the second list. He immediately had an immediate sense of "...something being wrong," to use his words. This recall pattern and AL's own reports convincingly portray visual imagery recall—he spontaneously "sees" the numbers in a space before him. The image was static, allowing for free access on demand. The images themselves have a strong mnemonic quality, increasing his recall capacity far above the norm.

8.4.1.2 Experiment AL.2: All 26 Alphabetical Characters

This next recall task involved all 26 alphabetical characters in an effort to further stretch AL's recall capabilities. He was first presented with the characters in a 2 x 13 columnar orientation, similar for that used in task one. The elements of the columns were again randomly chosen, and were presented in the following order: Col 1: F, L, M, U, S, O, W, T, A, C, D, J, H; Col 2: E, P, G, K, Z, N, R, Q, B, X, I, Y, V. AL requested to be allowed time to examine the list until he felt that he had it in his memory, and took 150 seconds to learn the list. His recall was no longer perfect in this task. He recalled the first column perfectly and rapidly, but recalled the letter

"F" for the "K" in the second column and omitted the "X". AL was aware that he was making errors. He reported that "...something was wrong early in the second row," and commented "I think there's one missing," when recalling the list at the point that he should have been reporting the "X".

The second part of the task consisted of recalling all 26 characters arranged in a 13 x 2 row layout. In this task, the rows were presented as: **Row 1:** R, D, H, Q, G, E, N, W, I, F, T, U, K; **Row 2:** Y, P, L, A, M, Z, V, C, S, O, B, X, J. AL took 120 seconds to study these rows, and then affirmed that he had the list memorised. He recalled all elements quickly and perfectly accurate. When asked to describe why this list was easier to recall than the columnar list, he ascribed it to a "...good pattern of memorising them." He then went on to describe his mnemonics:

> *The order that you have them in didn't spell words, but I could associate them with things very easily. I picture the colours before me, but then use the associations. The associations bring them all to the surface, and then I just read the items across. It's like it is written on a plank of lead, that sinks underwater, but the associations can bring it up to the surface, where I can read the items off. It starts to sink again immediately, but I can pull it up again with the association and read across while it's above the surface.*

AL could also read the rows backwards and pair-wise, as if the list was 13 columns of two letters. He could also begin recall in the middle of a row at a requested letter and proceed either direction through the rest of the elements. AL's reporting of his recall experience, his ability to randomly access elements in the list, and his manipulation of item order all support the assertion that he is "seeing" as some constant image and not merely a fanciful association.

8.4.1.3 Further memory tasks

AL took part in several more similar memory tasks with more elements by introducing numeric digits as stimuli. He succeeded in memorising all 26 alphabetical characters and 10 numeric digits presented as two columns of 18 elements after a three minute exposure. His only error was one of inversion of two characters, which persisted regardless of how he was requested to recall the elements. The errors that occurred when presented with a variety of item layouts continued to be those of simple inversion and omissions. Almost all omissions were accompanied by AL's assertions that he was certain that items were missing.

AL demonstrated higher proficiency with the memorisation of digits, and learned a list of 4 x 5 single digits in less than a minute, recalled very rapidly. His retention for individual elements from several of the lists was excellent: those lists that he reported as strong images were recalled days later, most frequently without prompting. AL began recalling elements from numeric lists quicker than elements from alphabetical character lists. His recall of elements from numerical lists was usually without error. AL used a variety of techniques in his mnemonics. The use of synaesthetic imagery underlies other processes, and provides a foundation for his recall and chunking methods. AL is an undergraduate psychology student, and has a cursory knowledge of chunking from an overview in memory. He describes chunking synaesthetic images into three items or digits for easier recall. Thus, he increases his memory capabilities with "normal" methods *and* synaesthetic percepts.

8.4.2 Examination of phonological and synaesthetic confusability in AL's short-term memory

Conrad & Hull (1964) demonstrated that recall performance for six-letter strings degraded when the component letters of the string were phonologically confusable (*e.g.*, "BCD"). One possible explanation for this effect is the similarity between elements generating difficulties in encoding in a short-term memory storage area, such as the *articulatory loop* proposed by Baddeley (1984). A memory test which uses phonologically confusable and dissimilar items as stimuli can provide important insights into the memory capabilities or deficits in subjects: Bub, Black, Howell, and Kertesz (1987) did just this in their assessments of an aphasic patient.

AL's colour-imagery is evoked regardless of the mode of presentation for numbers or letters, and this imagery is eidetic and mnemonic. Visual imagery clearly plays an important role in AL's memory. The question remains as to exactly how this imagery is encoded into AL's memory, and what limitations using synaesthetic imagery generates. He reports difficulties recalling sequences that contain repeated letters or digits, or short repeated letter and digit strings. This difficulty may arise from visual similarity in the resulting imagery pattern—repeated digits generate repeated colours in his "bar-code". This form of interference may be an effect similar to phonological confusibility: *synaesthetic confusibility* in encoding.

This synaesthetic confusibility could be demonstrated in several ways. Potentially different levels of interference could be experienced by AL for oral and visual presentational modes. That is, AL may experience more phonological interference for orally presented sequences, more synaesthetic interference for visually presented sequences, a combination of interference with different weightings for each mode, or perhaps only synaesthetic interference in one or both modes.

8.4.3 Experiment AL.3: Orally presented letter sequences examining phonological and synaesthetic confusability.

8.4.3.1 Aims and Introduction—Six letter strings

This experiment was developed with AL to examine potential interference in encoding that he might experience, as outlined in the previous section. To this end, he was orally presented with a variety of alphabetical letter sequences that varied in their constructive vocabularies, chosen for potential interference characteristics. Due to time constraints, only a subset of the full set of matching experiments was performed with visually presented stimuli.

8.4.3.2 Method

Sequences were assembled from three letter vocabularies, the first two (phonetically similar and dissimilar) chosen from a subset of the original experiment vocabularies (Conrad & Hull, 1964) and a vocabulary chosen for AL's personal correspondences. The personal vocabularies are composed of *colour similar* letters; that is, sets of colours with similar colour correspondences, grouped by these colours. These vocabularies were:

1. Phonetically similar (colour dissimilar): B, C, D, G, P, T, V
2. Phonetically dissimilar (colour dissimilar): F, K, Q, R, W, X, Z
3. Colour Similar
    - Blacks........................ H, L, P, W
    - Reds.......................... J, M, N, R
    - Browns.......................B, Q, T, X
    - Greys........................ C, U, V, Y

Each of the stimulus strings was composed of six characters chosen from a condition vocabulary. Since the *colour similar* vocabularies were each composed of only four characters, it was impossible to create stimulus strings of six characters without repeating characters. For this reason, *colour similar* condition strings were made up of characters from *two colour similar* vocabularies: the first three characters of the string drawn from one *colour similar* vocabulary, and the second three from another (*e.g.*, "RNMVCY"—"RNM" from the *red colours* and "VCY" from the *grey colours*).

8.4.3.3 Procedure

While sitting with the subject, the experimenter selected a set of stimulus strings for an experimental condition. The selection was performed out of the sight of the subject. A stimulus string was selected by randomly choosing character cards without replacement from the condition vocabulary, writing down the chosen characters, and then reshuffling the vocabulary characters cards. This process continued until all ten (10) stimulus strings were chosen for each experimental condition. The strings were presented in full sets for each condition. The order of conditions was

1. Phonetically similar;
2. Phonetically dissimilar;
3. Colour similar.

AL was presented with the sequences orally, with the elements of each stimulus string presented at the rate of one per second. At the conclusion of each presented string, AL was asked to reproduce the list, which was recorded for later analysis. AL's introspections during the experiment were also recorded, as they could provide potential insight into his mnemonic processes.

8.4.3.4 Results and discussion

Results were scored in a similar manner to the original experiment. All incorrect characters recalled were counted as a individual error. Thus, transposed characters were counted as two errors. Omitted characters would also counted as errors, although in the first three conditions no characters were omitted in AL's recalled sequences. Table 8.2 provides a summary of the results for the different conditions: the total errors made, percent of characters incorrectly recalled, and total number of error-free trials.

Orally Presented Letter Strings	*Mean Errors per trial*	*Characters in Error*	*Error Free Trials*
Phonetically Similar	2.3	38.0 %	2
Phonetically Dissimilar	0.8	13.3 %	6
Colour Similar	0.5	8.3 %	7

Table 8.2. Summary result statistics for different conditions in phonological confusability experiment.

These results suggest that AL experienced the same type of interference in the encoding of phonologically similar letters as normal individuals. No evidence is found for synaesthetic similarity generating more interference in encoding than phonologically dissimilar items.

There are two possible reasons for not finding synaesthetic interference. First, AL still has the phonologic information available for memory encoding: if he did not, he should not have exhibited interference between the phonologically similar and dissimilar items. The two sources of information with different weightings may together contribute to encoding into memory. The other explanation is that synaesthetic and phonetic encoding may have differing importance under different presentational modes. When listening to letters, phonetic information may be of greater importance; when reading, synaesthetic information may be more important. In addition to these factors, the experiment may have had a ceiling effect due to the relatively short length of the stimulus strings.

8.4.4 Experiment AL.4: Orally presented eight letter phonologically confusable strings.

8.4.4.1 Aims and Introduction

AL's performance with the six letter colour-similar strings was very proficient, and *evidence from the previous experiments demonstrate that he was capable of regularly* recalling much longer (*e.g.*, 10 element) sequences. Longer eight letter colour-similar strings were next used to examine how performance might change with greater demands on memory. These stimulus strings were composed in a similar manner to the six letter strings, with the first four elements from one colour-similar vocabulary and the latter four from another. The experiment was performed in an identical manner to the previous experiment (Experiment AL.3; see Section 8.4.3). A total of ten eight letter strings were presented for recall.

8.4.4.2 Results and discussion

AL's performance with eight letter strings was greatly degraded when compared to the recall for six-letter colour-similar strings. AL now made an average of 4.8 errors per string, with 60% of the characters being recalled in error, compared to 0.5 errors per string and 8.3% for the six letter strings. Only one of the trials for the eight letter condition was now error-free compared to seven in the six letter condition. Certainly a large portion of his change in performance must be ascribed to the increased length of the stimulus string; however, AL had already demonstrated his proficiency at much longer strings presented both aurally and visually. There were some interesting aspects to his recall performance that may provide some insight into his processing. While the letters were being read to him, AL looked up in space to his right, presumably towards his described area of visualisation. During recall, he would tend to gesture in the air, and reported that he was pointing to the elements that he was trying to visualise. Additionally, he also omitted entire segments rather than attempting guesses, which was never a pattern in his responses before. AL reported that he had far less difficulty with the red and pink hued letters ("R, N, M, J") than the remaining letters, as the hues were very distinct; the greys and blacks giving him quite a bit of difficulty. On three trials he omitted entire segments. Table 8.3 shows the stimuli and responses for those three trials.

Trial	*Stimuli*	*Response*
2	QTBXVYUC	WQX————
5	XTQBYUCV	XQT————
9	YVCULHWP	————

Table 8.3. Recalled strings with missing string segments for eight-letter condition.

In each of these examples, it was the grey letter sequence which was not recalled; in Trial 9, the black letter sequence was also forgotten. In light of these suggestive results, and AL's own introspection into his performance, AL may have been experiencing synaesthetic confusibility in encoding.

8.4.5 Experiment AL.5: Visually presented seven letter phonologically confusable strings.

8.4.5.1 Aims and Introduction

AL was tested by visually presenting strings from the phonologically confusable vocabularies in order to test further differences in memory performance that might be due to mode of presentation. This was done in a manner similar to the experiment of Bub et al (1987). For this task, the individual letters were placed in the centre of a 3 x 5" index cards in a 72-point Helvetica font. The letters were shown to AL for one second each, moving the cards from the front to the back of the pack. After the last letter was exposed, AL began recalling as much of the sequence as possible. The character strings were randomised by shuffling the pack of cards between trials, and the experimenter recorded each pattern before exposure began. In this experiment, the character strings were each composed of the total set of seven letters from the phonologically confusable category, without repetition of any letter in the string. Seven letter strings were chosen as stimuli due to the maximum possible length of the vocabulary and restriction of non-repetition of any element.

8.4.5.2 Results and discussion

AL's performance for the visually presented seven letter strings was superior in all regards to his performance for the aurally presented six letter sequences. He committed a mean of 1.7 errors per trial compared to 2.3; 24% of all characters were recalled in error compared to 38%; and, 4 trials were error-free compared to 2. Only one character was omitted in recall, and that was the sole error for that trial.

These results provide evidence that the visual component of the strings facilitates more accurate memory performance. Whereas a non-synaesthetic subject may internally vocalise the characters in an effort to retain them in short-term memory, AL did not need to rely upon this technique as much as non-synaesthetes. Instead, he could employ the colour-imagery directly for a mnemonic aid. The input of the auditory system directly into such a processing area like the articulatory loop may necessitate its use and thus its interference with visual imagery. Since the articulatory loop was not required for the visual presentations, the synaesthetic imagery could be utilised more effectively.

8.5 Conclusions

The performance of the synaesthetes tested in cross-modal experiments suggests that the fixedness of synaesthetic associations hinders the use already associated of the imagery qualities as SoCs. That is, if colours are already associated with some quality of sound, that association keeps the synaesthete from using colour to represent evaluate qualities. It is difficult for synaesthetes to say how closely representations approximate their imagery: all representations that do not adequately represent their imagery are equally incorrect.

DS's performance further demonstrates that synaesthetic percepts impinge upon attention, as discussed in the previous chapter. DS's percepts actively distracted her from the task, keeping her from looking at representations that were in conflict with her imagery until that imagery ceased. Finally, the conflict between the two representations generated a strong emotional response due to the "incorrect" nature *of the representation. This again demonstrates that* synaesthetes are fixed in their associations, and cannot understand other associations.

The idiosyncratic nature of each synaesthetes experiences make standardised synaesthetic tests virtually impossible. This necessitates customizing experiments in a case-study approach for examining synaesthetic performance. The paucity of memory experiments with synaesthetes has led to a reliance on anecdotal reports of increased mnemonic abilities, while performance may not lie outside of normal ranges. AL's dramatic STM abilities are evident when performing digit-span experiments, but not so apparent in his WMS score. Additionally, those reported cases of increased memory are exceptional (Subject AL and Luria's patient S), and should not be generalised to all synaesthetes.

Synaesthetic associations provide additional information that may facilitate improved memory abilities. Some of the factors that aid memory are the specific qualities of the synaesthete's imagery. AL's colour-associations remain constant when numbers are combined into larger strings, thus requiring only a one-to-one mapping of colour onto symbol. Subjects like AJ experience different sensations when characters are combined together, which necessitates more base-level correspondences between symbol combinations and synaesthetic percepts. Accordingly, AL incorporates his synaesthetic percepts into his recall strategies, and has better than average STM abilities; AJ's associations are pleasant for her, but provide her no advantage in STM.

AL's ability to recall letters in arbitrary orders specified by the experimenter further demonstrates that his imagery is spatial, and that he recalls from that spatial representation. He is recalling by actively picturing the letters simultaneously. Some of the errors he commits are similar to those that would be made by someone losing their place while performing the same task with a printed list of items.

AL's synaesthetic imagery does not replace all of the processes that would normally be involved with numbers and letters. Experiment AL.3 (see Section 8.4.3) demonstrates that he still experiences interference in encoding generated by phonologically confusable letters, indicating that AL does use sound as well as imagery. However, his performance when given longer strings orally (Experiment AL.4; see Section 8.4.4) suggests that he does use visual imagery in the encoding. Memory performance with visual presentation of phonologically confusable letters (Experiment AL.5, see Section 8.4.5) demonstrates that different modes of presentation change the importance that synaesthetic imagery plays in the memory task.

Much more quantitative research with synaesthetes is required. The findings that have been reported in the last twenty years provide only the most rudimentary foundation upon which to build synaesthetic theories. After the psychological community agrees upon what form of perception constitutes a synaesthetic condition, more progress may be made in determining incidence of synaesthesia.

Such information will provide opportunities for more experiments into the common advantages and deficits that synaesthesia affords its possessors.

9 Modified Stroop Experiments With Synaesthetes

9.1 Overview

This chapter will examine the types of interference that colour-number synaesthetes may experience when presented with information that is incongruent to their personal coloured associations. Two experiments will be presented that examine several synaesthete's performances in modified Stroop experiments. Experiment One was conducted with two synaesthetic subjects, and consisted of the three standard Stroop tasks and additional conditions of numbers coloured congruently and incongruently to the synaesthete's idiosyncratic associations. The synaesthetes experienced significant interference for the standard Stroop task and the incongruently coloured numbers. Experiment Two was another Stroop task with three synaesthetes and three controls, with individual presentation of stimulus elements and responses provided by a voice interface into a computer program. Stroop interference and incongruently coloured number interference was again observed.

9.2 Introduction

The opportunity of working with a subject over an extended time period allowed for the creation of experiments tailored to an individual synaesthete and their specific synaesthetic imagery. For example, the synaesthete could more precisely describe their imagery for specific associations with the use of a colour computer monitor and computer painting software package. This produced a set of stimuli that could be used with slightly modified well-established psychological experiments.

The experiments in this chapter set out to examine the depth at which associated colour percepts are experienced by colour-number synaesthetes. If they occur at an early processing level, are evoked involuntarily, and impinge upon the synaesthetes attention, as hypothesised, it will be possible to observe interference from these percepts in a controlled experiment. Melara (1989a) determined that integral dimensions of auditory and visual information can generate cross-modal interference, similar to Garner Interference (see Section 1.2). This chapter examines the cross-modal interference that is experienced by the fixed cross-modal perceivers: synaesthetes.

9.3 Experiment Fourteen: Coloured-numeric Stroop Interference

9.3.1 Aims and Introduction

One of the most robust effects in experimental psychology is the Stroop colour–word interference task (Stroop, 1935). Stroop attributed his findings of delay in the naming of coloured ink in incongruently coloured colour names to a high-level interference generated by the incongruent stimuli. The semantic content inherent in the written word interfered with the content generated by identifying the ink colour and *verbalising it*. *Stroop accepted the* explanation of Peterson, Lanier, and Walker (1925; cited in MacLeod, 1991) that the process of reading a colour-name word involved a single response while the process of identifying an ink colour evoked many processes of which the identification of the object's specified attribute was just one. In addition to the interference findings, Stroop also observed that performance on the incongruent colour-ink naming task improved significantly with practice.

Several modifications to this paradigm may be necessary to account for interference experienced from coloured-language synaesthesiae. Asserting that reading a colour-word evokes a *single* response is over-simplified: certainly this would be the case for a true coloured-lexical synaesthete. The very elements constituting the word—the letters themselves—may evoke several associations in addition to, and potentially in conflict with, the content of the word itself.

As outlined in Chapter Seven, synaesthetic associations are fixed and present as far back as the synaesthete can recall. This level of deep association of colour with language and numbers may generate interference when presented in incongruently coloured letters, words, and/or digits. A modified Stroop task using incongruently coloured stimuli for an individual synaesthete should reveal higher levels of interference, measured as increased delay times between presentation of stimulus and subject response.

One immediate problem that presents itself is the difficulty in predicting how the two types of interference might interact: would the interference from semantic information be so strong as to mask additional interference from the synaesthetic component, would they work together in a additive fashion, or would a more complex interaction of interference occur? A simplification of stimuli to simple single character stimulus items facilitates examination of synaesthetic interference before proceeding to a traditional Stroop task with colour names. The use of a non-colour word as a stimulus may have different levels of associations for different types of synaesthetes. That is, some experience individual sensations from each component letter (AL), others a colour sense from the overall word (AJ), and still others different associations for individually presented letters and complete words. The use of a single alphabetical character raises another potential problem, as associations exist between the letter and colour name ("G" equated with "green").

The use of coloured-numeric stimuli is proposed for the investigation of synaesthetic Stroop-like interference. This reduces the amount of semantic information that may be present in the stimulus. That is, normals should not have colour qualities associated with digits, whereas the synaesthetes have these associations by definition. All of the coloured-language synaesthetes available for experiments also had coloured-numerals, but not all synaesthetes experienced either coloured letters, coloured phonemes, or coloured words. One final reason for using digits is the smaller number of single-digit numerals compared to the letters in the English alphabet.

This experiment included all of the basic Stroop experiment trials (Spreen & Strauss, 1991). These tasks provide baseline measurements of colour-naming, incongruent non-colour words, and incongruently coloured colour-names that can be compared to population norms.

9.3.2 Method

9.3.2.1 Subjects

Two synaesthetes who have colored-numbers as one facet of their synaesthesia took part in the experiment: Subject AL and Subject MG (see Chapters Seven and Eight for details).

9.3.2.2 Equipment

A NeXT TurboColor workstation was used to determine subject colour associations as well as presenting the coloured stimuli for the naming task. Response times were manually recorded by the experimenter with a digital stopwatch.

9.3.2.3 Stimuli

9.3.2.3.1 "Cards"

Each task used a computer generated "card"—a computer image of an A4 sized sheet. Each card had five rows of four separate stimulus items making a total of twenty stimulus items in all. This is a replication of the standard Stroop test for neuropsychological assessment (Spreen & Strauss, 1991). Figure 9.1 shows a partial set of twelve stimuli used for AL in the coloured-circles condition.

9.3.2.3.2 Colours

Because each synaesthete had different colour associations, the colours in each of the five conditions were different for each synaesthete, but constant between trials for that particular synaesthete. The process of deciding upon colour and number subsets to be used in the experiment is described in the procedure section. Hereafter, the individual set of four colours decided upon between synaesthete and experimenter will be referred to as the *digit colour set.*

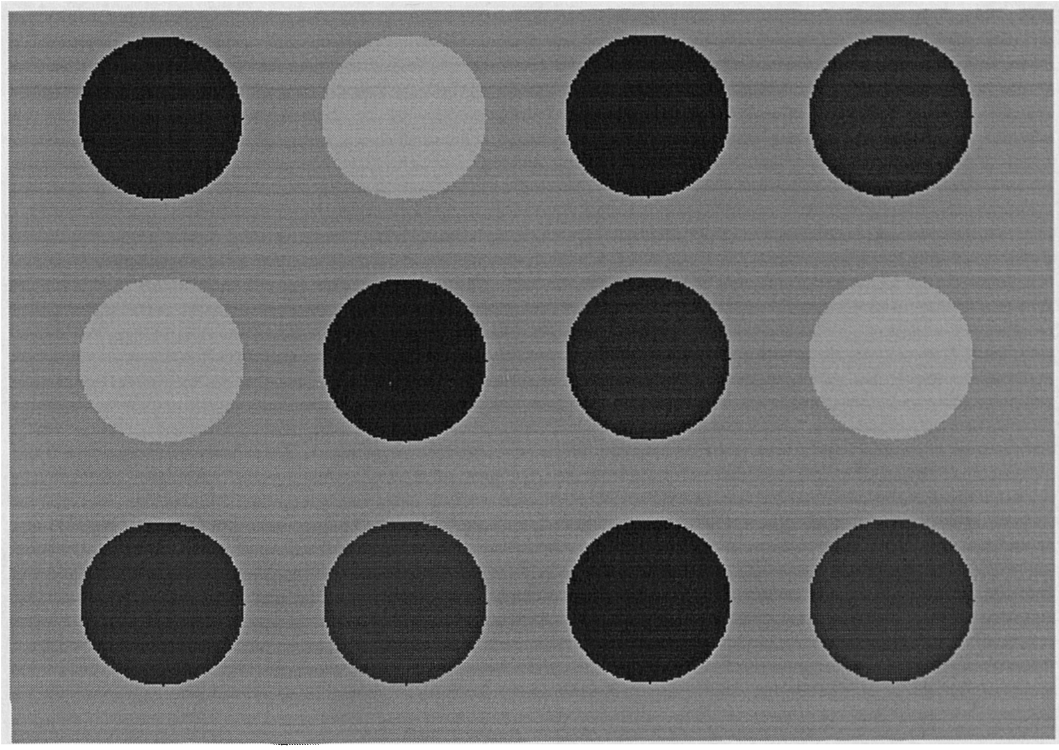


Figure 9.1. Sample coloured stimuli for AL used for the Coloured-circles condition in Experiment One.

9.3.2.3.3 Coloured Numbers

Each synaesthete had different digits that met the criteria necessary for the procedure, necessitating the use of different digits for each synaesthete. Digits were always selected from the range of 2 to 9. Digits were displayed with the Helvetica font on the NeXT VDU for all synaesthetes. This font was chosen due to all subjects describing visualising the characters as sans-serif. When shown some different fonts, each chose Helvetica as an adequate representation of their visualisations. Figure 9.2 shows a sample subset of coloured numeric stimuli used for AL in the congruently-coloured trials.

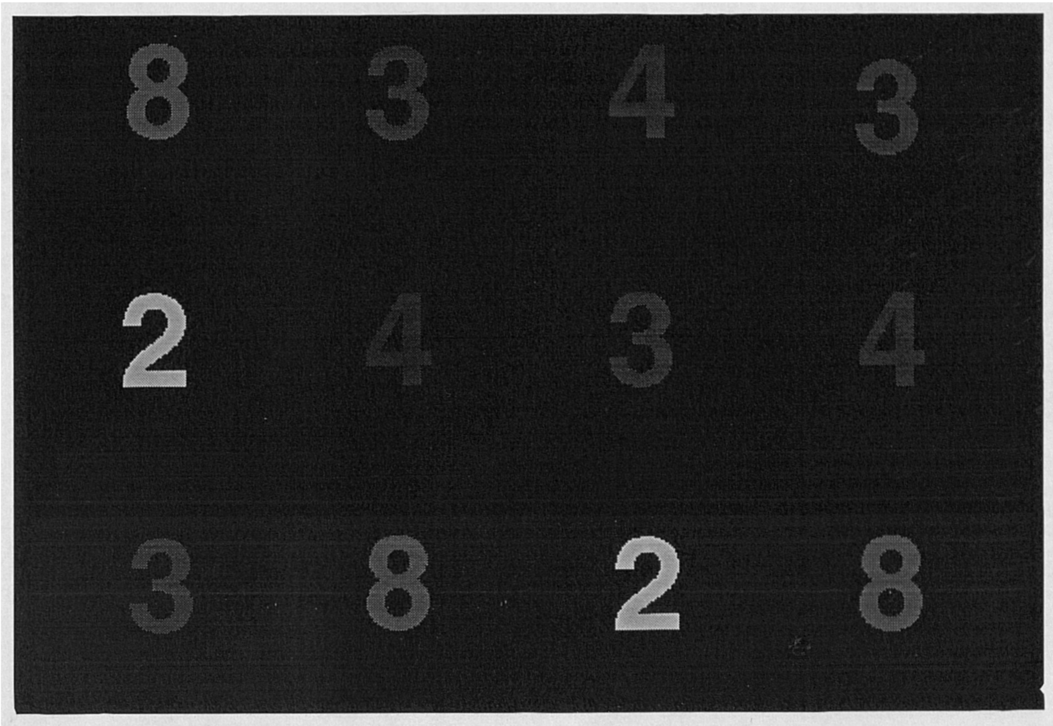


Figure 9.2. Sample coloured-numeric stimuli for AL used in the congruently-coloured trials.

9.3.3 Procedure

For the synaesthetic subjects, the experimenter worked with the synaesthete to identify four single digit numbers that were most simply coloured. Some of the subjects report that the visual qualities of numbers vary in colouration or shading along the number's shape (*e.g.*, *1* for subject AL was one colour and surrounded by an aura of another colour). For this reason, it was important to select digits that were composed of a single solid hue.

Another criterion for selecting digits was the ease in naming the perceived colour. Priority was given to those correspondences that could be identified by a basic colour category word, such as "blue" rather than "azure". From the synaesthetes set of ten single digits, four were chosen to match the four colours from the standard Stroop task. The experimenter and synaesthete worked together to generate as close of a match for each coloured number as possible. When the synaesthete was satisfied that the computer representation of the digits closely matched his/her visualisations, the experiment proper commenced.

The synaesthete was then presented with five colour naming tasks. These were performed in a block of trials. Different subjects performed the five tasks in different orders. The five colour naming tasks were:

1. *Coloured-spots (CS)*: stimuli consisted simply of circles coloured using the digit colour set;
2. *Non-colour words (NCW)*: stimuli consisted of the non-colour words from the standard Stroop experiment coloured using the digit colour set ("when", "hard", "and", and "over");
3. *Incongruently coloured words (ICW)*: stimuli consisted of the colour names for the digit colour set incongruently coloured (*e.g.*, the word "Red" presented in a green hue);
4. *Congruently coloured numbers (CN)*: stimuli consisted of the digits from the digit colour set correctly coloured for that synaesthete, and;
5. *Incongruently coloured numbers (IN)*: stimuli consisted of the digits from the digit colour set incorrectly coloured.

In each task, the synaesthete was instructed to respond with only the colour of the stimulus element. Subjects were instructed to correct their own errors before proceeding to the next element. The experimenter kept separate track of uncorrected errors. The total time to name each element on each "card" was recorded.

9.3.4 Results

9.3.4.1 Main Stroop Effect

Table 9.1 shows the mean RTs and standard deviations for the two subjects in all conditions. Synaesthetes exhibit the same general robust Stroop effect for incongruently coloured words, as well as the controls. Subject AL took part in seven complete blocks of trials, three on the first day and four more on the following day. Subject MG was only available for one day of testing, and took part in four blocks of trials. Figure 9.3 shows this same data in a box and whisker plot.

Subject		CS	NCW	*ICW*	*CN*	*ICN*
MG	*Mean*	10.93	12.71	21.51	11.64	13.99
MG	*SD*	1.14	0.99	1.15	0.96	0.94
AL	*Mean*	8.40	9.39	14.02	10.59	12.23
AL	*SD*	0.82	1.06	1.11	1.01	0.89

Table 9.1. Means and standard deviations for MG and AL in all five Stroop tasks (CS = Coloured Spot; NCW = Non-colour word; ICW = Incongruently-coloured word; CN = Congruently coloured number; IN = Incongruently coloured number).

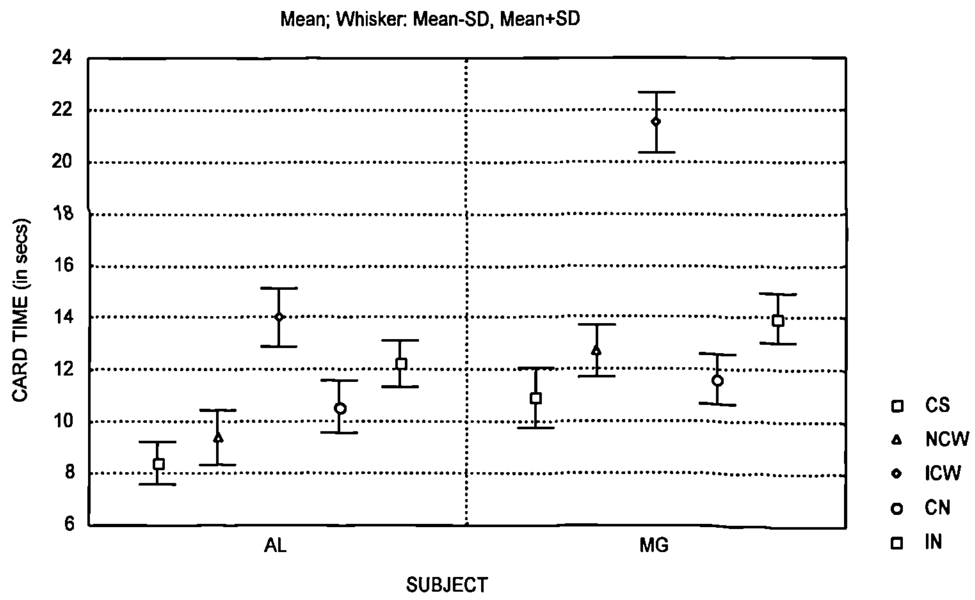


Figure 9.3. Means and standard deviations for card reading times in each experimental condition, by subject (CS = Coloured Spot; NCW = Non-colour word; ICW = Incongruently-coloured word; CN = Congruently coloured number; IN = Incongruently coloured number).

Many of the conditions reveal significant differences in mean card reading time for different conditions. Table 9.2 shows the results of separate *t*-tests run on the means from the different conditions and their corresponding *p*-values for Subject AL. Those *t*-scores with significant *p*-values are marked as: less than .05 with an asterisk (*); less than .01 with a dagger (†); and less than .001 with a double dagger (‡). As seen in this table, significant differences in reading times are observed with the naming colour of coloured spots (CS) faster than all other experimental conditions. Incongruent-colour word RTs (ICW) are significantly slower than the non-colour word times (NCW), demonstrating the classic Stroop effect. ICW RTs are also significantly slower than congruently-coloured numbers RTs (CN). Incongruent-colour number naming (IN) is significantly slower than NCW times, and slower than congruent-colour number times (CN). Figure 9.4 plots the individual times for each of AL's card-reading times on each trial.

Subject AL	*CS*	*NCW*	*ICW*	*CN*	*IN*
Colour Spot (CS)	—	−2.591*	−9.480‡	−4.864†	−10.676‡
Non-Colour Word (NCW)	2.591*	—	−9.093‡	−2.149	−7.286‡
Incongruent Colour Word	9.480‡	9.093‡	—	6.186‡	2.909*
Congruent Number (CN)	4.864†	2.149	−6.186‡	—	−2.951*
Incongruent Number (IN)	10.676‡	7.286‡	−2.909*	2.951*	—

Table 9.2. Individual *t*-test scores for Subject AL RTs in the different Stroop conditions in Experiment One (* : $p < .05$; † : $p < .01$; ‡ : $p < .001$; a *t*-test computed with six degrees of freedom).

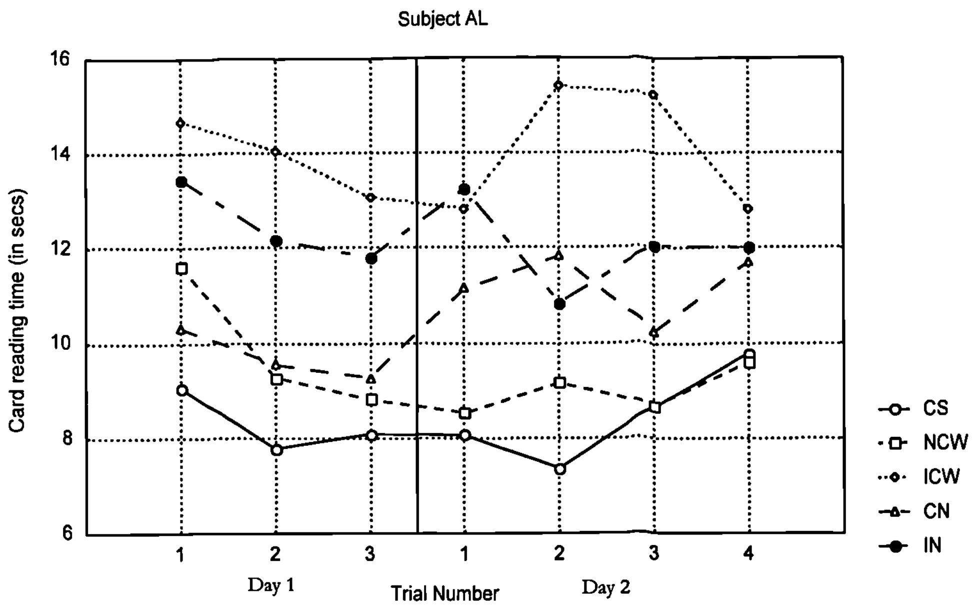


Figure 9.4. Subject AL's individual scores for each of the card reading trials on separate days (CS = Coloured Spot; NCW = Non-colour word; ICW = Incongruently-coloured word; CN = Congruently coloured number; IN = Incongruently coloured number).

Table 9.3 shows the *t*-scores with significant *p*-values marked for Subject MG, with the data in the same arrangement as in Table 9.2. As seen in this table, significant differences in reading times are observed with the naming colour of coloured spots (CS) faster than incongruent-colour words (ICW) and incongruent-colour numbers (IN). The main Stroop effect is again observed with ICW times significantly slower than the non-colour words times (NCW). ICW times are also significantly slower than and congruent-colour number times (CN). IN times are significantly slower than the naming times of NCW and CN. Figure 9.5 plots the individual times for each of MG's card-reading times on each trial.

Subject MG	*CS*	*NCW*	*ICW*	*CN*	*IN*
Colour Spot (CS)	—	– 1.918	– 65.729‡	– 2.708	– 7.433†
Non-Colour Word (NCW)	1.918	—	– 10.382†	1.262	– 2.159
Incongruent Colour Word	65.729‡	10.382†	—	32.521‡	25.785‡
Congruent Number (CN)	2.708	– 1.262	– 32.521‡	—	– 5.031*
Incongruent Number (IN)	7.433†	2.159	– 25.785‡	5.031*	—

Table 9.3. Individual *t*-test scores for Subject MG RTs in the different Stroop conditions in Experiment One (* : $p < .05$; † : $p < .01$; ‡ : $p < .001$; a *t*-test computed with three degrees of freedom).

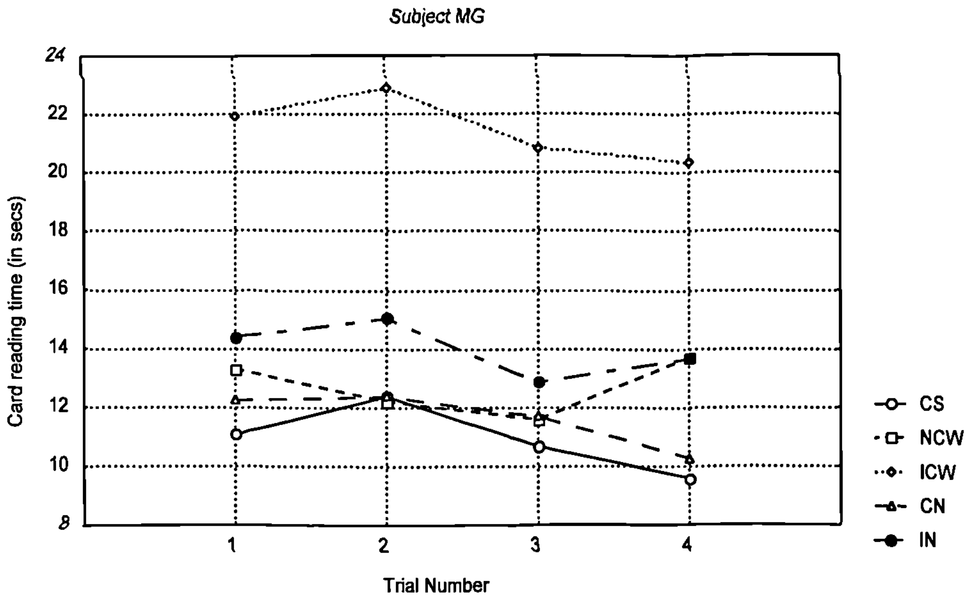


Figure 9.5. Subject MG's individual scores for each of the card reading trials on separate days (CS = Coloured Spot; NCW = Non-colour word; ICW = Incongruently-coloured word; CN = Congruently coloured number; IN = Incongruently coloured number).

9.3.5 Discussion

Both synaesthetic subjects strongly demonstrate the classic Stroop effect. This indicates that interference is experienced from the semantic content of the ICW in the process of colour naming.

For both subjects, non-colour word (NCW) times were not significantly different from congruent number (CN) times. The synaesthetic correspondences between colour and number could not then be said to be facilitating colour response. However, both of the subjects have strong colour associations with each letter in the alphabet, so that interference could be present in the NCW and ICW conditions. Subject MG reports single-colour associations with each of the four NCW words ("when"—bluey-grey [sic]; "hard"—brown; "and"—white; "over"—white with a tinge of yellow). Her responses may have this synaesthetically interfering association changing her RTs.

In addition to the classic Stroop effect the evidence indicates that interference is experienced between the two coloured number conditions, CN and IN. This supports the hypothesis that colour association from a digit are automatically evoked in the process of reading numbers. Additionally, this also fulfils the criteria from Chapters Two and Seven: that the imagery is involuntary and constant in nature. More importantly, the synaesthetic images are processed in a manner that interferes with colour naming. This lends support to the hypothesis that the associations are deep, and that they impinge on the synaesthete's attention.

9.4 Experiment Fifteen: Coloured-numeric Stroop using single presentations of stimuli

9.4.1 Aims and Introduction

The Stroop effect observed in Experiment One and in the classic Stroop experiment all use the response times for a set of similar stimuli (*e.g.*, twenty items on one card). One potential difficulty with the results is that all stimulus items are visible during the task; that is, the subject may look ahead to the next element or set of elements before naming. A method that presented an individual stimulus in a random order would not permit the subject to use secondary response cues from the digits.

The practical problem of accurately recording RTs arose with this change to single item presentation. A computer program that detects the subject's voice response to the stimulus and record the delay in response time was developed to solve this problem. The interface was sensitive only to the volume level of the voice, and recorded response times in milliseconds between onset of the stimulus and the beginning of the subject's response.

For this experiment, non-colour words ("when", "hard", "and", and "over": the *NCW* condition from Experiment One) were not used. Instead, a congruently-coloured word condition was used to examine if the congruency of colour might help facilitate quicker response times. This might be an important factor in determining response time, as the single stimulus presentation might reveal an additional effect: the change of congruency between different trials.

9.4.2 Subjects

Three coloured-number synaesthetes (one male, AL; two females, MG & LH), and three non-synaesthetes (one male, CONHH; two females, CONTC & CONRA) were used as controls. The controls were all members of the CCCN (*Centre for Cognitive & Computational Neuropsychology*) at Stirling University. Each control subject was paired with a synaesthete, matched by gender and approximately matched for age, and used the same stimulus set as that synaesthete: CONHH with AL, CONTC with MG, and CONRA with LH. Subjects were not naïve to the concept of synaesthetic associations between colour and number; however, none were told in advance of the experiment's specific aims. They were also unaware of the "correct" colour and digit congruent pairing.

9.4.3 Method

9.4.3.1 Equipment

A NeXT TurboColor computer was used to create and display visual stimuli. In addition, the computer was programmed to record verbal responses via a microphone. The RT was recorded as the delay between the onset of the visual stimulus and the first occurrence of a sound level over an experimenter-set audio threshold. Bad subject responses, such as an unintentional noise or incorrect colour name, were noted by the experimenter.

9.4.3.2 Stimuli

9.4.3.2.1 Coloured Patches

Coloured rectangles of the same width and height of the largest word in the subjects stimulus set were presented in solid colour. The colours for the rectangles were taken from the colour-digit set for the synaesthete or control.

9.4.3.2.2 Coloured Words and Numbers

The words of the four colours chosen by the synaesthete were displayed in congruent and incongruent colour. For Subjects AL and CONHH, these words were "Yellow", "Red", "Green", and "Brown", corresponding to the digits 2, 3, 4, & 8, respectively; Subjects MG and CONTC were presented with the words "White", "Yellow", "Green", and "Blue", corresponding to the 1, 2, 4, and 6, respectively. Subjects LH and CONRA were presented with the words "Green", "Red", "Yellow", and "Blue", corresponding to the digits 3, 4, 6, & 7 respectively.

Five experimental conditions were created from the coloured patches, words, and numbers. The conditions were:

1. *Solid Coloured Rectangles (CR)*
2. *Incongruently Coloured Words (ICW)*
3. *Congruently Coloured Words (CW)*
4. *Incongruently Coloured Numbers (ICN)*
5. *Congruently Coloured Numbers (CN)*

Each condition had four different elements (four different words, digits, or rectangles). Each single stimulus was presented five times, for a total of 100 stimulus items composing a single block.

9.4.4 Procedure

The colours and digits that were to be used as stimuli were first determined by the same process outlined in the procedure section of Experiment One (Section 9.3.3). This process preceded the trials for the synaesthetic subjects; the controls skipped this selection process, as the colours were predetermined. For all subjects, the experimenter adjusted the voice response threshold so that a response from the subject in a normal speaking voice was detected by the computer. At this point, the experiment proper commenced.

Subjects were seated 50 centimetres from the computer VDU. They were instructed to respond with the colour of the presented stimulus, and to disregard the content of word and digit. They were told to make their responses verbally in a loud, clear voice.

The experiment began with the presentation of one of the stimulus items. The entire stimulus set was presented in a random order, with no two consecutive trials from the same condition. The item remained visible on the screen until the subject responded. After the response, the screen went blank for three seconds, and then was followed by the next stimulus. If the subject made an incorrect response, the experimenter pressed the space bar on the computer's keyboard during the three second pause. This keystroke was recorded in the data file for identification in analysis.

The subject continued to make responses in this manner until the entire stimulus set had been presented. At the completion of the last trial, a "Thank You" message was displayed on the screen. After a short break, the subject proceeded to the next block of trials. After completing the two blocks, the subject was informed of the experiment's aims and debriefed.

9.4.5 Results

The mean RTs with accompanying standard deviations, and standard errors are presented in a box plot in Figure 9.6 for the synaesthetic subjects, and Figure 9.7 for the controls.

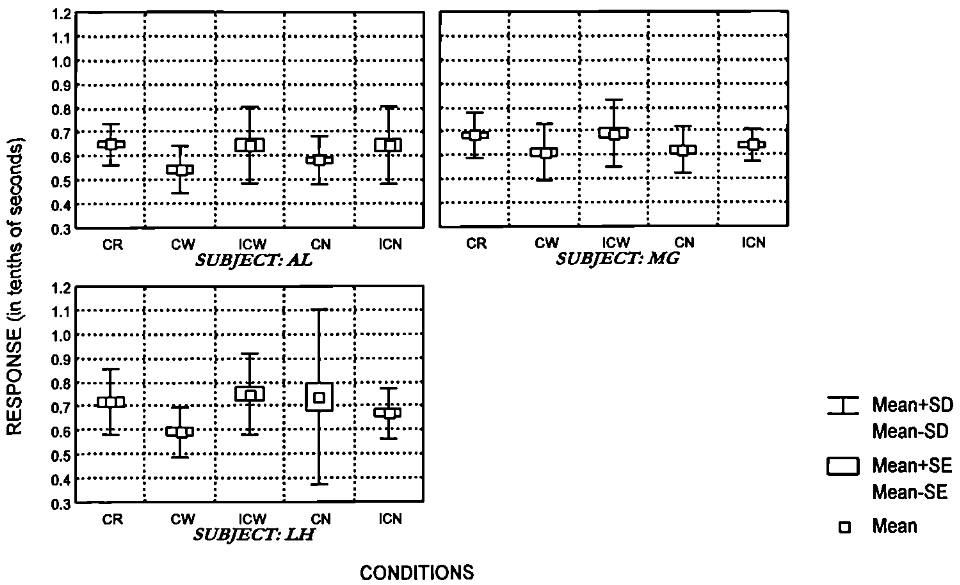


Figure 9.6. Means, standard deviations, and standard errors for each of the three synaesthetic subjects in Experiment Fifteen. The abscissa is divided into the separate conditions (CR = Coloured rectangle; CW = Congruent colour word; ICW = Incongruently colour word; CN = Congruently coloured number; IN = Incongruently coloured number).

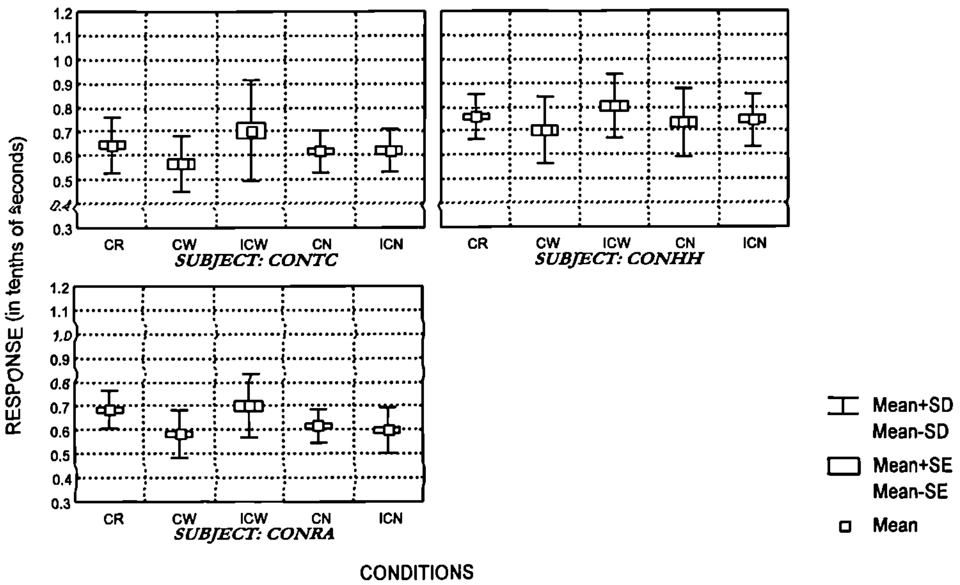


Figure 9.7. Means, standard deviations, and standard errors for each of the three control subjects in Experiment Fifteen. The abscissa is divided into the separate conditions (CR = Coloured rectangle; CW = Congruent colour word; ICW = Incongruently colour word; CN = Congruently coloured number; IN = Incongruently coloured number).

Baseline colour naming times in the coloured rectangle condition were nearly identical for the synaesthetic group (mean = 0.68; sd = 0.11) and controls (mean = 0.70; sd = 0.11). Table 9.4 presents the *t*-scores for each subject's RTs between different pairs of conditions. Significant interference in colour naming for incongruent colour words compared to congruent colour words is observed for all subjects, both synaesthetes and controls, demonstrating the main Stroop effect. Additionally, all subjects demonstrate significant facilitation in RTs for congruently coloured items over the solidly coloured patches. No subjects demonstrate significant interference in RTs between coloured rectangles and incongruently-coloured words. Only one synaesthetic subject, AL, demonstrates significant interference between congruently coloured and incongruently coloured numbers. None of the control subjects demonstrate any interference between the digit conditions.

SUBJECTS	*CR vs. CW*	*CR vs. ICW*	*CW vs. ICW*	*CN vs. ICN*
AL	$5.132^{\ddagger}$	0.153	$-3.432^{\ddagger}$	-2.067^{*}
MG	$2.922^{\dagger}$	-1.262	$-2.658^{\dagger}$	-1.070
LH	$4.599^{\ddagger}$	-0.905	$-5.002^{\ddagger}$	1.172
Control1	$2.993^{\dagger}$	0.954	$-3.693^{\ddagger}$	-0.262
Control2	2.188^{*}	0.765	$-3.325^{\dagger}$	-0.295
Control3	$5.129^{\ddagger}$	-0.647	$-4.539^{\ddagger}$	0.858

Table 9.4. Individual *t*-test scores ($t(78)$) for all subjects' performance in the different Stroop conditions for Experiment Fifteen ($*$: $p < .05$; † : $p < .01$; ‡ : $p < .001$). (CR = Coloured rectangle; CW = Congruent colour word; ICW = Incongruently colour word; CN = Congruently coloured number; IN = Incongruently coloured number).

Subject AL and Subject MG demonstrate faster response times for congruently coloured numbers compared to solid colour rectangles (AL: $t(78) = 3.199, p < .002$; MG: $t(78) = 2.848, p < .01$).

9.4.6 Discussion

The method of using single item stimulus presentation still demonstrated the main Stroop effect for all subjects, with RTs increased for incongruently coloured words (ICW) compared to congruently coloured words (CW). This further demonstrates the robustness of the Stroop effect. All subjects also had quicker RTs for the CW trials in comparison to the coloured rectangle (CR) conditions. This facilitation effect in the CW conditions is due to the content of the colour word, in the same manner that the interference is generated in the ICW conditions. Although instructed to ignore the content of the word, the processing of the written word is involuntary, and the semantic content either interferes with or aids in the colour naming.

Subject AL experienced interference from the incongruently coloured numbers (ICN) when compared to congruently coloured numbers (CN) as hypothesised. Neither of the other two synaesthetes showed such interference. Also, Subject AL and MG both experienced significant facilitation from the CN stimuli.

One possible explanation for these results lies in the nature of the numeric stimuli in comparison to the words, and the varying strength of a synaesthete's associations. The nature of the single presentation naming task may make it easier for some subjects to ignore the synaesthetic content of the numbers. That is, it is possible that those synaesthetes with stronger associations also find those associations more attention demanding. Subject LH did describe that her colour-number associations as a "sense of a colour", while Subjects AL and MG could more clearly describe the imagery that they experienced. In addition, Subject LH had only colour-number associations, while these associations were only part of a larger set of synaesthetic experiences for AL and MG.

It is also possible that the single item presentation method may introduce new some problems in the attempt to eliminate the opportunity for subjects to look forward at the full stimulus set. One problem with the method is that the content of a stimulus item changes on almost every trial. The constant oscillation between facilitating and interfering content may actually reduce effects, by making it easier for subjects to ignore the content of the stimuli, as they are instructed.

9.5 Conclusions

Experiment Fourteen shows that the synaesthetic subjects with coloured-number and coloured-letter synaesthesia exhibit the classic Stroop effect with increased response times for incongruently coloured colour-names. In addition, these synaesthetes exhibit a synaesthetic Stroop-like effect when presented with numbers that are coloured incongruently for that subject. These results show that these synaesthetic subjects experience interference from both the semantic content of a word and the synaesthetic content of a number.

Experiment Fifteen used a new method of presenting single stimuli with voice response to eliminate a potential "look-ahead" effect present when subjects can see the entire set of stimuli when responding. A control group was used to examine non-synaesthetes colour-naming response times. The method was adequate to reveal the classic Stroop effect, as well as a facilitating effect for congruently coloured colour-names. The modified synaesthetic Stroop effect was exhibited by one subject. Congruently coloured numbers also facilitated response times in two synaesthetes. The varying strength of associations for different synaesthetes was hypothesised as a factor that dictates how much interference is generated by a number.

The use of a modified Stroop task with synaesthetes as subjects demonstrates that synaesthetic associations are involuntarily evoked, and can interfere with other cognitive processes. This interference results in part from the attention-demanding nature of synaesthetic imagery. Further experiments with a larger sample of synaesthetes need to be performed to determine how much strength of synaesthetic associations change the levels of interference, and to examine if non-synaesthetic subjects who learn to pair colour with items also experience interference.

10 Conclusions

The experiments in the preceding chapters have each demonstrated how of using *scales of comparison* (SoCs) can be used to explain both cross-modal perception and synaesthetic perception. The SoC hypothesis suggests that when presented with a novel stimulus, individuals can judge that stimulus on another sensory scale for the purposes of evaluation. That evaluation is performed so that the encountered stimulus may be used by the perceiver in a meaningful fashion.

In the process of creating a cross-modal matching task, the experimenter sets up the boundaries of a scenario wherein comparisons are made using different SoCs from different sensory modalities. The task presented requires the subject to make a measurement of a stimulus in one modality and evaluate it on a perceptual scale in the other modality. The process is the same as the subject uses when encountering a novel stimulus, but the choice of SoCs with which to evaluate the stimulus with is restricted. In a non-laboratory setting, the subject could evaluate the novel stimulus on any salient scale she might choose.

The experiments in Chapter Three revealed that individual variation in cross-modal matching experiments may be present but obscured by sample trends. To explain this in terms of the SoC hypothesis, individuals can align the same SoCs in different directions, or use different SoCs and produce consistent matching performance across modalities. Even when presented with auditory stimuli and visual stimuli that each vary in one dimension only, subjects demonstrate both directions of alignment between SoCs when forming cross-modal matches. The finding that subjects can match stimuli consistently, but perform these matches using different SoCs and alignments is to be expected if the SoC hypothesis explains some of the origins of synaesthesia. Synaesthetic associations are idiosyncratic; the ability for subjects to vary in their use of even rudimentary SoCs would certainly generate differences when subjects persist in using those SoCs (and thus form permanent synaesthesia).

It is still important to examine what the majority of subjects use as SoCs and alignments for those SoCs in cross-modal matching. Primarily this consistency across subjects indicates that the SoCs are cognitively proximal for many subjects, and thus have qualities in common. These commonly used SoCs and alignments show up as main sample effects in cross-modal matching results. Experiment One used a simple two-alternative forced choice design to demonstrate that although a strong effect could be observed for matching increasing pitch height with increasing visual brightness, this was not the only regular pattern of matching subjects chose. Opposite direction of alignment between SoCs yields consistent matching performance, and suggests that different subjects could be using individual criteria for evaluating the stimuli. It might appear that subjects match the visual SoC of decreasing visual brightness with increasing pitch height, while they could instead be using the visual SoC of increasing contrast. The important result from the experiment is that subjects map from one scale onto another lawfully, and use this evaluation to form their matches.

Experiments Two and Three each took the simple two-alternative forced choice design from Experiment One and increased the complexity of the auditory stimuli, making them more musical in content. Main sample effects were still observed in the cross-modal matching, and this again provides evidence pointing to those qualities of the music that were most salient to the majority of subjects. The observed effects of specifically musical qualities acting as SoCs for cross-modal comparisons (*e.g.*, consonance) cannot be explained by a suprasensory representation for visual and auditory brightness. In terms of an SoC hypothesis, these effects can be accounted for *by subjects using scales of musical* qualities as SoCs for performing evaluations. As the stimuli increased in complexity, the opportunity for subjects to use several different qualities of the stimuli as SoCs also increased. Individual analyses of matching performances indicates that different subjects used a variety of scales in different alignments.

Experiments Four, Five, and Six in Chapter Four extended the results of the first three experiments by varying the hue of the visual stimuli. In these experiments, the visual stimuli were now much more complex, as they varied in a quality that cannot be easily ordered: *colour*. With the opportunity for even more variation in SoCs that subjects could use, sample effects decreased. In Experiment Five, no sample effects were observable, although the individual analyses still demonstrated that consistent matching was being performed by subjects. Presumably those subjects that did not show consistent matching in the analyses were still performing in an internally consistent manner. However, if they were using SoCs that did not correlate with the experimenter-hypothesised scales, this performance is hidden. Ultimately, this is the limitation of all cross-modal experiments. If the experimenter restricts the boundaries of the cross-modal task adequately, he can demonstrate regular matching because only a few scales are available for subjects to use. When the task becomes more unbounded by making stimuli more complex, only those SoCs used by many subjects and anticipated by the experimenter will be observable. Other consistent matching is obscured, and perhaps even regarded as spurious.

Experiment Seven, Eight, and Nine in Chapter Five each continued to extend the cross-modal results, but focused on potential musical properties of an auditory stimulus that could be matched cross-modally with vision. Experiment Seven demonstrated that subjects could use specifically learned musical qualities of auditory information as an effective SoC. Experiments Eight and Nine each show that subjects were more sensitive to the timbre of a musical note than the pitch height when each were allowed to vary. This further suggests that the suprasensory representation of brightness is too simplistic to explain the varied performances in cross-modal matching experiments.

In speculative terms, visual information in conjunction with cross-modal matching techniques might be used in a instructional setting to aid in learning fundamentals of music. Subjects may not have been aware of any consonance difference present in dyad or chords; however, regular matching performances reveal that they were not only capable of distinguishing such a difference, but also able to use this discrimination in another task. Even if they could not express this consonance difference between dyads in words, they were demonstrably sensitive to it. The cross-modal matching task eavesdrops, or indirectly measures the subjects' internal representations of music. Using cross-modal evaluation techniques that allow students to access these internal representations indirectly might be used instructionally to help learn the concept of consonance. Possibilities exist for the use of "visual metaphors" to teach other fundamental musical concepts.

Each of the previous experiments used stimuli that were fundamental in their component qualities, although the stimuli varied along several psychological dimensions. The experiments in Chapter Six endeavoured to demonstrate that the SoC hypothesis could still explain matching performed with scales of a much higher-order. Affective and descriptive scales represented by language (adjectives) could be used consistently by subjects for the evaluation of melodic fragments as well as entire musical pieces. Individuals use the same SoCs in the same alignments because they have the added information from language to create and use these SoCs.

Some scales, such as a *Happy/Sad* scale, are used to represent a continuum of evaluative qualities while the underlying measured qualities are quite different in nature. Experiment Ten showed that subjects could use such a scale for evaluating a piece of music. Experiment Eleven demonstrated that the quality of *happiness* is communicated in short fragments while *sadness* takes longer to establish. Experiments Twelve and Thirteen each demonstrated that subjects can use multiple SoCs from language and even landscape pictures for matching. One limitation of this type of experiment is that the more complex the SoCs become, the more difficult it is to deduce what subjects are actually using to make their comparisons.

Chapter Seven showed how common qualities of categorical synaesthesiae can be accounted for by using a model wherein synaesthesia develops through a process of individuals using SoCs when encountering new stimuli that require evaluation. When encountering a stimulus from a new conceptual group, a measurement of the stimulus is taken. With no stable set of measurements from a conceptual group to compare the stimulus with, another *existing* set of measurements is borrowed for the comparison. If this comparison is adequate to perform the evaluation, the individual may persist in using the SoC until the scale becomes an integral part of perceiving the stimuli.

The fixedness of synaesthetic associations and categorical nature of these associations naturally arises from this SoC development model. So also does the fact that the imagery does not degrade in potency when repeatedly evoked: it is simply an integral part of the perceptual process. The three-dimensional qualities of the imagery for visual synaesthetes arises from the common use of spatial qualities in vision as an effective SoC.

Chapter Eight further demonstrated the fixedness in synaesthetic percepts by examining the performances of synaesthetes when participating in the cross-modal matching experiments from the earlier chapters. The nature of the matching experiment restricts the available SoCs that subjects can use. When a subject already has cross-modal associations that use available stimulus dimensions, they cannot separate those dimensions out of their joined percepts and use them as SoCs. They can perform comparisons on other SoCs which may be different from those used by normal subjects, and thus it is difficult for an experimenter to determine what SoCs they are using. Nevertheless, their performance in cross-modal tasks can still be consistent and use the same SoC mechanism, but appear to be random.

Experiments with coloured-number synaesthete AL provided evidence that the consistent use of synaesthetic imagery can aid in memory tasks. The consistent and flexible nature of his associations in conjunction with his superior short-term memory abilities suggests that the lawful use of a SoC may aid in some mental processes. This also provides some direct evidence for the reputed increased mnemonic powers of some synaesthetes. AL's performance when presented with phonologically confusable strings of letters demonstrated that mode of presentation changes the saliency of his synaesthetic imagery for a particular task. Phonological properties of words were still used by AL for encoding into memory, even in the presence of the accompanying visual imagery. Visual presentation of those same strings removed the phonologic component of the stimuli, and AL relied more upon his synaesthetic imagery. Thus, synaesthetic imagery is an integral part of his perception that affords him additional context-dependent abilities.

Chapter Nine further explored the depth at which synaesthetic associations exist as well as their involuntary nature. Modifications on the standard Stroop task created experiments that demonstrated interference from synaesthetic percepts. This provides more evidence that the imagery is an integral part of perception for synaesthetes, and directly impinges on their attention. One limitation of these experiments came from the approximations of each synaesthete's imagery. The synaesthetes each had difficulties adequately explaining how to physically produce their imagery, and the quality of any computer-produced approximation may be insufficient to generate more pronounced interference. Stroop task experiments using more accurate representations could reveal if interference is a function of the stimulus's representational quality.

The concept of using scales of comparison to make evaluations can be useful for the analysis of intra-modal comparisons as well as cross-modal matching. As a developmental model, it can help explain the formation of schemata. Additionally, it may be possible to determine which scales of comparison can be directly presented to aid in the development of a new concept. Children's toys and primers regularly use colours to learn numbers and letters and both the Yamaha system and Tobin Music programme each use coloured notes for teaching music.

An effort should be made to use scales of comparison in a consistent manner to aid in instruction. It may be that the variety of colour systems and shapes used for teaching concepts are used in such a varied manner that they can rarely be used as a consistent SoC. If children were always presented with the same use of instructional SoCs, perhaps categorical synaesthetic perception would be more common and useful.

References

ALEXANDER, K. R. & SHANSKY, M. S. 1976. Influence of hue, value, and chroma on the perceived heaviness of colors. *Perception & Psychophysics*, **19**, 70–74.

ANDREWS, D. P. 1964. Error-correcting perceptual mechanisms. *Quarterly Journal of Experimental Psychology*, **16**, 104–115.

ARISTOTLE. 1966. *The Works of Aristotle.* Edited by W. D. Ross, Oxford: Clarendon Press.

AUERBACH, C. & SPERLING, P. 1974. A common auditory-visual space: Evidence for its reality. *Perception & Psychophysics*, **16 (1)**, 129–135.

BADDELEY, A. D. 1984. Memory theory and memory therapy. In B. Wilson and N. Moffat (eds.), *The Clinical Management of Memory Problems*, 5–27. London: Croom Helm.

BARON-COHEN, S. 1996. Is there a normal phase of synaesthesia in development? *PSYCHE: An Interdisciplinary Journal of Research on Consciousness*, **2** (27).

BARON-COHEN, S., HARRISON, J., GOLDSTEIN, L. H. & WYKE, M. 1993. Coloured speech perception: Is synaesthesia what happens when modularity breaks down? *Perception*, **22**, 419–426.

BARON-COHEN, S., HARRISON, J., LOADER, & RAHMAN. 1993. Unpublished thesis. Cambridge University.

BARON-COHEN, S., SMITH-LAITTAN, F., HARRISON, J., & BOLTON, P. 1996. Synaesthesia: prevalence and familiarity. *Perception*, **25 (9)**, 1073—1079.

BERLIN, B. & KAY, P. 1969. *Basic Colour Terms: Their Universality and Evolution.* Berkeley: University of California Press.

BERNARD, J. W. 1986. Messaien's Synaesthesia: The correspondence between color and sound structure in his music. *Music Perception*, 4 (1), 41–68.

BERNSTEIN, I. H., & EDELSTEIN, B. A. 1971. Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, **87** (2), 241—247.

BINET, A. Étude sur un nouveau cas d'audition colorée. *Revue Philosophique.* **33**, 461—464.

BIRREN, F. 1950. People who think in colors. *Science Digest*, **28**, 1–5.

BLACK, M. & KLIG, S. 1970. Dimensional learning across sensory modalities in nursery school children. *Journal of Experimental Child Psychology*, 9, 166–173.

BLEULER, E. & LEHMANN, K. 1881. *Zwangmassige Lichtempfindungen durch Schall und verwandte Erscheinungen.* Leipzig: Fues Verlag.

BORING, E. G. & STEVENS, S. S. 1936. The nature of tonal brightness. *Proceedings of the National Academy of Sciences*, **22**, 514–521.

BRUNER, J. S. & KOSLOWSKI, B. 1972. Visually preadapted constituents of manipulatory action. *Perception*, **1**, 3–14.

BUB, D., BLACK, S., HOWELL, J., & KERTESZ, A. 1987. Speech Output Processes and Reading. In C. Coltheart, G. Sartori, and R. Job (eds.), *The Cognitive Neuropsychology of Language*, 79–110. London: Lawrence Erlbaum Associates.

CALKINS, M. W. 1895. Synaesthesia. *American Journal of Psychology*, **7**, 90–107.

CHACKSFIELD, K. M., BINNS, P. A., & ROBINS, V. M. 1973. *Music & Language with Young Children.* Wallop, Hampshire: BAS Printers Limited.

COHEN, N. E. 1934. Equivalence of brightness across modalities. *American Journal of Psychology*, **46**, 117—119.

CONRAD, R. & HULL, A. J. 1964. Information, acoustic confusion and memory span. *British Journal of Psychology*, **55**, 429–432.

COOKE, D. 1959. *The Language of Music.* London, Ely House: Oxford University Press.

CYTOWIC, R. E. 1996. *The Neurological Side of Neuropsychology.* MIT Press: Cambridge, Massachusetts.

CYTOWIC, R. E. 1995. Synesthesia: Phenomenology and Neuropsychology. *Psyche*, **2** (10).

CYTOWIC, R. E. 1993. *The Man Who Tasted Shapes.* Abacus: London.

CYTOWIC, R. E. 1989. *Synesthesia: a union of the senses.* New York: Springer-Verlag.

CYTOWIC, R. E. & WOOD, F. B. 1982a. Synesthesia: I. A review of theories and their brain basis. *Brain and Cognition*, **1**, 29.

CYTOWIC, R. E. & WOOD, F. B. 1982b. Synesthesia: II. Psychophysical relations in the synesthesia of geometrically shaped taste and colored hearing. *Brain and Cognition*, **1**, 36–49.

DEUTSCH, D. 1986. A musical paradox. *Music Perception*, **3**, 275–280.

DOWLING, W. J. & HARWOOD, D. L. 1986. *Music Cognition.* London, England: Academic Press Limited.

DYER, F. N. 1973. Interference and facilitation for color naming with separate bilateral presentations of the word and color. *Journal of Experimental Psychology*, **99** (3), 317–317.

ELLSON, D. G. 1941a. Hallucinations produced by sensory conditioning. *Journal of Experimental Psychology*, **28**, 1–20.

ELLSON, D. G. 1941b. Experimental extinction of an hallucination produced by sensory conditioning. *Journal of Experimental Psychology*, **28**, 350–361.

FLOURNOY. 1893. *Des phâenomâenes de synopsie (audition colorâee): photismes, schâemes visuels, personnifications.* Fâelix Alcan: Paris.

GALILEO GALILEI. 1623. *The achievement of Galileo. Edited with notes by James Brophy and Henry Paolucci.* Twayne Publishers: New York [1962].

GALTON, F. 1883. *Inquiries into Human Faculty and its Development.* Everyman: London.

GARNER, W. R. 1974. Attention: The processing of multiple sources of information. In E. C. Carterette & M. P. Friedman (Eds.) *Handbook of perception* (Vol. 2, pp. 23–59). Academic Press: New York.

GARNER, W. R., & FELFOLDY, G. L. 1970. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, **1**, 225—241.

GLICKSOHN, J., SALINGER, O., & ROYCHMAN, A. 1992. An exploratory study of syncretic experience: Eidetics, synaesthesia and absorption. *Perception*, **21**, 637–642.

GOUK, P. 1988. The harmonic roots of Newtonian science. In J. Fauvel, R. Flood, M. Shortland, and R. Wilson (eds.), *Let Newton be! A new perspective on his life and works*, 100–123. Oxford, England: Oxford University Press.

GUIRAO & STEVENS. 1964. Measurement of auditory density. *Journal of the Acoustical Society of America*, **36**, 1176—1182.

HARRISON, J. & BARON-COHEN, S. 1995. Synaesthesia: reconciling the subjective with the objective. *Endeavour*, **19** (4), 157–160.

HART, T. H. 1909. Report of a case of psychochromesthesia. *The Journal of Abnormal Psychology*, **6**, 157–160.

HARTSHORNE, C. 1934. *The philosophy and psychology of sensation.* Kennikat Press: Port Washington, New York.

HELMHOLTZ, H. 1954. *On the sensations of tone as a physiological basis for the theory of music.* Unabridged and unaltered publication of the original 2nd edition of the Ellis translation of *Die Lehre von den Tonempfindungen* and published originally by Longmans & Co., Dover Publications, Inc., 180 Varick Street, New York.

HORNBOSTEL, E. M. VON. 1931. Über Geruchschelligkeit. *Pflügers Archiv für die Gesamte Physiologie*, **227**, 517—538.

HOUSTON, R. A. 1917. Newton and the colours of the spectrum. *Science Progress*, **12**, 250–264.

JAMES, J. 1993. *The Music of the Spheres: Music, Science and the Natural Order of the Universe.* London, England: Abacus.

JORDAN, D. 1917. The colour of letters. *Science*, **1187**, 311—312.

KANT, I. 1781. *Critique of pure reason.* Translated by Norman Kemp Smith. Humanities Press: New York [1950].

KARWOSKI, T. F., & ODBERT, H. S. 1938. Color-Music. *Psychological Monographs.* **50** (2).

KARWOSKI, T. F., ODBERT, H. S. & OSGOOD, C. E. 1942. Studies in synaesthetic thinking. II. The role of form in visual responses to music. *Journal of General Psychology*, **26**, 199–222.

KÖHLER, W. 1947. *Gestalt psychology, an introduction to new concepts in modern psychology.* Liveright Pub. Corp.: New York.

KROHN, W. D. 1892. Pseudo-chromesthesia, or the association of colors with words, letters, and sounds. *American Journal of Psychology*, **5**, 20–41.

LANGER, J. & ROSENBERG, R. G. 1975. Symbolic meaning and color naming. *Journal of Personality and Social Psychology*, **4** (4), 364–373.

LANGFELD, H. S. 1926. Synesthesia. *Psychological Bulletin*, **23**, 599–602.

LEWKOWICZ, D. J. & TURKEWITZ, G. 1980. Cross-modal equivalence in early infancy: auditory-visual intensity matching. *Developmental Psychology*, **16**, 597—607.

LOCKE, J. 1690. *An Essay Concerning Human Understanding: Book 3.* Basset: London.

LOCKHEAD, G. R. 1972. Processing dimensional stimuli: a note. *Psychological Review*, **79** (5), 410—419.

LOCKHEAD, G. R. 1979. Holistic versus analytic process models: a reply. *Journal of Experimental Psychology: Human Perception*, **5** (4), 746—755.

LUNDBORG, H. 1923. Die Vererbung des "Favenhörens" und gewisser anderer psychisches Eigenschaften innerhalt eines Schwedischen Geschlechts. *Acta med. Scand.*, **59**, 221–229.

LURIA, 1968. *The Mind of a Mnemonist.* New York: Basic Books.

MACLEOD, C. M. 1991. Half a century of research on the Stroop Effect: An integrative review. *Psychological Bulletin*, **109** (2), 163–203.

MACPHARLAIN, C. 1908. *Binneas Nam Bard (Bardic Melody): A book in which the poems, songs, and ditties of the Scottish Gaels are exhibited with their airs.* Vol. 1. Stirling, 43 Murray Place: Aonghas Mac Aoidh.

MARKS, L. E. 1974. On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology*, **87**, 173–188.

MARKS, L. E. 1975. On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*, **82**, 303–331.

MARKS, L. E. 1978. *The Unity of the Senses: Interrelations Among the Modalities.* New York: Academic Press.

MARKS, L. E. 1987. On cross-modal similarity: Perceiving temporal patterns by hearing, touch, and vision. *Perception & Psychophysics*, **42** (3), 250–256.

MARKS, L. E. 1989. On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, **15** (3), 586–602.

MARKS, L. E., HAMMEAL, R. J., & BORNSTEIN, M. H. 1987. Perceiving similarity and comprehending metaphor. *Monograph Social Research in Child Development*, **52** (1), 1—102.

MARKS, L. E., & SLAWSON, A. W. 1966. Direct test of the power function for loudness. *Science*, **154** (**752**), 1036—1037.

MARKS, L. E. & STEVENS, J. C. 1966. Individual brightness functions. *Perception & Psychophysics*, **1**, 17–24.

MAURER, D. 1993. Neonatal synesthesia: Implications for the processing of speech and faces. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, 109–124. Netherlands: Kluwer Academic Publishers.

MAURER, D., & MAURER, C. 1988. *The World of the Newborn*, Basic Books: New York.

MAURER, D. & MONDLOCH, C. J. 1996. Synesthesia: A stage of normal infancy? *Proceeding of the International Society for Psychophysics, Padua*, October.

MELARA, R. D. 1989a. Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, **15** (1), 69–79.

MELARA, R. D. 1989b. Similarity relations among synesthetic stimuli and their attributes. *Journal of Experimental Psychology: Human Perception and Performance*, **15** (2), 212–231.

MELARA, R. D. & MARKS, L. E. 1990a. Dimensional interactions in language processing: Investigating directions of levels of crosstalk. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16** (4), 539–554.

MELARA, R. D. & MARKS, L. E. 1990b. Processes underlying dimensional interactions: Correspondences between linguistic and nonlinguistic dimensions. *Memory and Cognition*, **18** (5), 477–495.

MELARA, R. D. & MARKS, L. E. 1990c. Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance*, **16** (2), 398–414.

MELARA, R. D., & O'BRIEN, T. P. 1987. Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, **116**, 323—336.

MELTZOFF, A. N. & BORTON, R. W. 1979. Intermodal matching by human neonates. *Nature*, **282**, 403–404.

MURATA, K., CRAMER, H., & BACH-Y-RITA, P. 1965. Neuronal convergence of noxious, acoustic, and visual stimuli in the visual cortex of the cat. *Journal of Neurophysiology*, **28** (6), 1223—1229.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. 1957. *The measurement of meaning*. University of Illinois.

PAUL, D. A. 1946. Color and line in music. *School Arts*, 45, 248.

PAULESU, E., HARRISON, J., BARON-COHEN, S., WATSON, J. D., GOLDSTEIN, L., HEATHER, J., FRACKOWIAK, R. S., & FRITH, C. D. 1995. The physiology of coloured hearing. A PET activation study of colour-word synaesthesia. *Brain*, **118** (pt 3), 661—676.

PEACOCK, K. 1985. Synesthetic perception: Alexander Scriabin's Color Hearing. *Music Perception*, **2** (4), 483–506.

PLATT, J. R. 1990. Perceived pitch class of isolated musical triads. *Journal of Experimental Psychology: Human Perception and Performance*, **16** (2), 415–428.

PLUMMER, E. & STARK, J. 1956. A picture painted in sound. *Etude*, **74**, 15, 50.

POLZELLA, D. J. & BIERS, D. W. 1987. Chromesthetic responses to music: Replication and extension. *Perceptual and Motor Skills*, **65**, 439–443.

POLZELLA, D. J. & KUNA, A. 1981. Chromesthetic responses to the music of G. F. Handel. In E. Kinger (Ed.), *Imagery: concepts, results, and applications.* Plenum Publishing Corp.: New York, N. Y.

POMERANTZ, J. R. 1985. *Selective attention: Differences between Stroop and Garner interference.* Paper presented at the 26th Annual Meeting of the Psychonomic Society, Boston, MA.

POMERANTZ, J. R., & GARNER, W. R. 1973. Stimulus configurations in selective attention tasks. *Perception & Psychophysics*, **14**, 565—569.

POULTON, E. C. 1989. *Bias in Quantifying Judgments.* East Sussex, U. K.: Lawrence Erlbaum Associates Ltd.

PRATT, C. C. 19. The spatial character of high and low tones. *Journal of Experimental Psychology*, **13**, 278–285.

RADER, C. M. & TELLEGEN, A. 1987. An investigation of synesthesia. *Journal of Personality and Social Psychology*, **52 (5)**, 981–987.

RAINES, T. H. 1909. Report of a case of psychochromesthesia. *The Journal of Abnormal Psychology*, **4**, 249–252.

RAKOWSKI, A. 1995. The perception of pitch and timbre. *Proceedings of Sympozjum Re}yserii I In}ynierii D}wi⊗ku, Warsaw*, 26–28.

RÉVÉSZ, G. 1923. Über audition colorée. *Zeitschrift für Angewandte Psychologie*, **21**, 308–332.

RICH, G. J. 1919. A study of tonal attributes. *American Journal of Psychology*, **30**, 121—164.

RIGGS, L. A. & KARWOSKI, T. 1934. Synaesthesia. *British Journal of Psychology*, **25**, 29–41.

RIZZO, M. & ESSLINGER, P. J. 1989. Colored hearing synesthesia: an investigation of neural factors. Neurology, 39 (6), 781—784.

SCHNEIDER, B. A. & BISSETT, R. J. 1981. The dimensions of tonal experience: A nonmetric multidimensional scaling approach. *Perception & Psychophysics*, 30 (1), 39–48.

SIMPSON, L. & MCKELLER, P. 1955. Types of synaesthesia. *Journal of Mental Science*, 101, 141–147.

SIMPSON, R. H., QUINN, M. & AUSUBEL, D. P. 1956. Synaesthesia in children: Association of colors with pure tone frequencies. *The Journal of Genetic Psychology*, 89, 95–103.

SINGLETARY, C. E. 1956. Color-music. *School Arts*, 56, 29–31.

SLONIMSKY, N. 1955. Musical oddities: Rimsky-Korsakov's association between colors and keys. *Etude*, 73, 4.

SMITH, A. B. 1918. Color Association. *Science*, 48, 395–396.

SPREEN, O. & STRAUSS, E. 1991. *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford: Oxford University Press.

STEVENS, S. S. 1934. The attributes of tones. *Proceedings of the National Academy of Sciences*, 20, 457–459.

STEVENS, S. S., & GUIRAO, M. 1963. Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology*, 66 (2), 177–188.

STEVENS, S. S., GUIRAO, M., & SLAWSON, A. W. 1965. Loudness, a product of volume times density. *Journal of Experimental Psychology*, 69, 503—510.

STRERI, A. 1987. Tactile discrimination of shape and intermodal transfer in 2- to 3-month-old infants. *British Journal of Developmental Psychology*, 5, 213—220.

STRERI, A., & PECHEUX, M. G. 1986. Tactual habituation and discrimination of form in infancy: a comparison with vision. *Child Development*, 57 (1), 100—104.

STROOP, J. R. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.

TAYLOR, I. K. 1963. Phonetic symbolism re-examined. *Psychological Bulletin*, 60 (2), 200–209.

VERNON, P. E. 1930. Synaesthesia in music. *Psyche*, 10, 22—40.

WALKER, A. R. 1978. Perception and music notation. *Psychology of Music*, 6, 21—46.

WALKER, A. R. 1981. The presence of internalised images of musical sounds. *Council for Research in Music Education*, 66–67, 107—112.

WALKER, A. R. 1987. The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perceptual Psychophysiology*, 42 (5), 491—502.

WALKER-ANDREWS, A. S. & GIBSON, E. J. 1986. 'What develops in bimodal perception?', in L. P. Lipsitt and R. Rovee-Collier (eds.), *Advances in Infancy Research (Vol. 4).* Ablex: Norwood, N. J.

WARREN, R. M. & POULTON, E. C. 1966. Lightness of grays: Effects of background reflectance. *Perception & Psychophysics*, 1, 145–149.

WATT, R. L. 1988. *Visual Processing: Computational, Psychophysical and Cognitive Research.* East Sussex, U. K.: Lawrence Earlbaum Associates Ltd.

WATT, R. L. & ASH, R. L. 1997. Experimental analysis of the perception of sequences of tones. *Proceedings of Third Triennial ESCOM Conference, Uppsala*, 611–612.

WATT, R. L., GERSTLEY, L. D., & ASH, R. L. 1996. Semantic Content in Wagner's "Siegfried". *In preparation.*

WHIPPLE, G. M. 1900. Two cases of synaesthesia. *American Journal of Psychology*, **11**, 377–404.

WICKER, F. W. 1968. Mapping the intersensory regions of perceptual space. *American Journal of Psychology*, **81**, 178–188.

WILLIAMS, J. M. 1976. Synaesthetic adjectives: A possible law of semantic change. *Language*, **52**, 461–478.

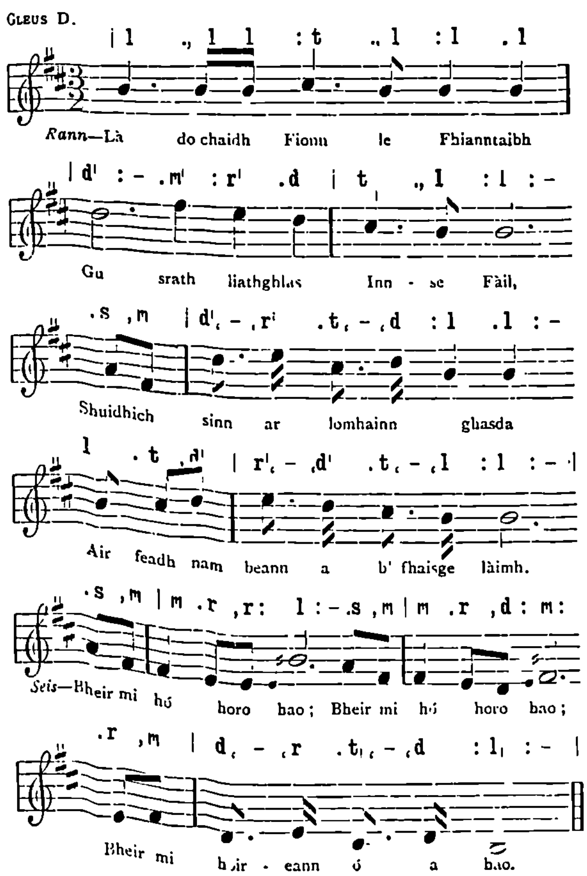
ZATORRE, R. J., EVANS, A. C., & MEYER, E. 1994. Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, **14** (**4**), 1908–1919.

Appendix A:
31 Unaccompanied Gaelic folktunes used as stimuli in Chapter Six

Tunes 1 & 2

(left and right, respectively)

Garabh is na Mnathan.

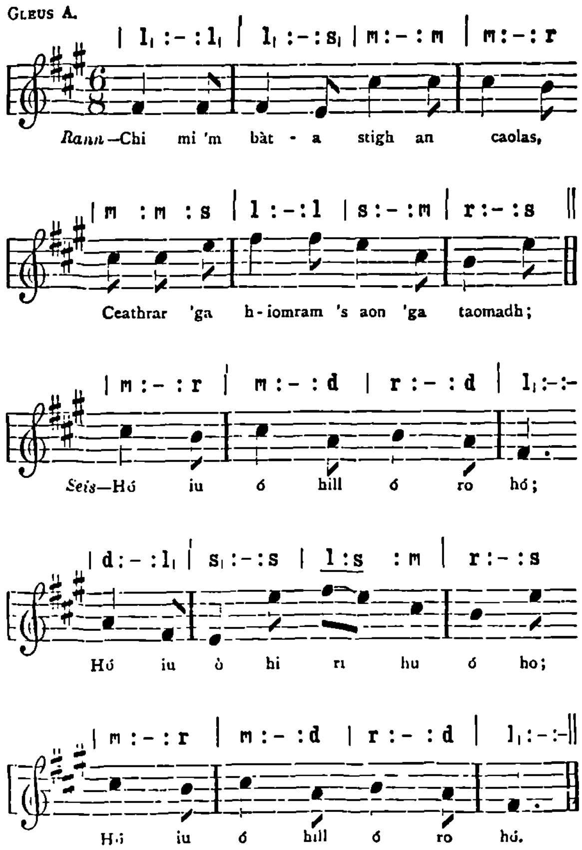


Laoidh an Amadain Mhòir.

Tunes 3 & 4

(left and right, respectively)

Cumba nan Iasgairean.



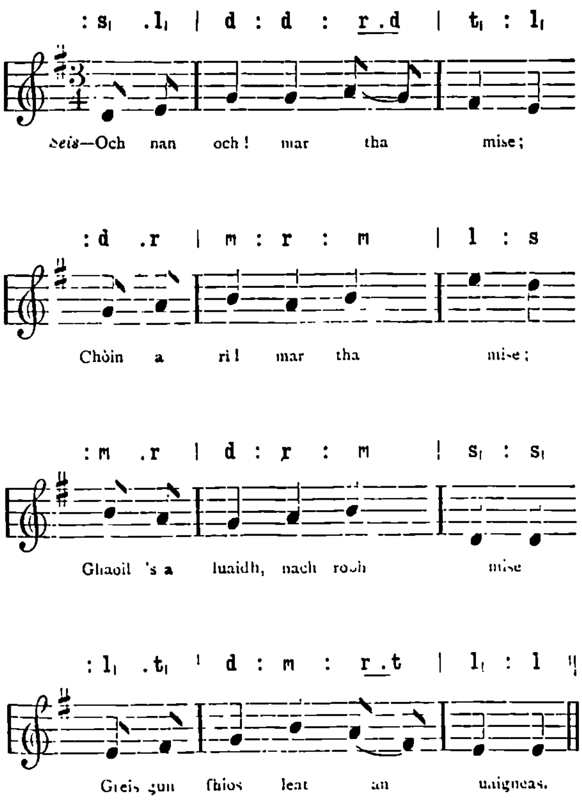
Cumha do Mhac Dhonnachaidh Ghlinne Faochain.

Tunes 5 & 6

(left and right, respectively)

Och nan Och!

Gleus G.



Cumha do Chòirneal Iain Camshron.

Gleus D

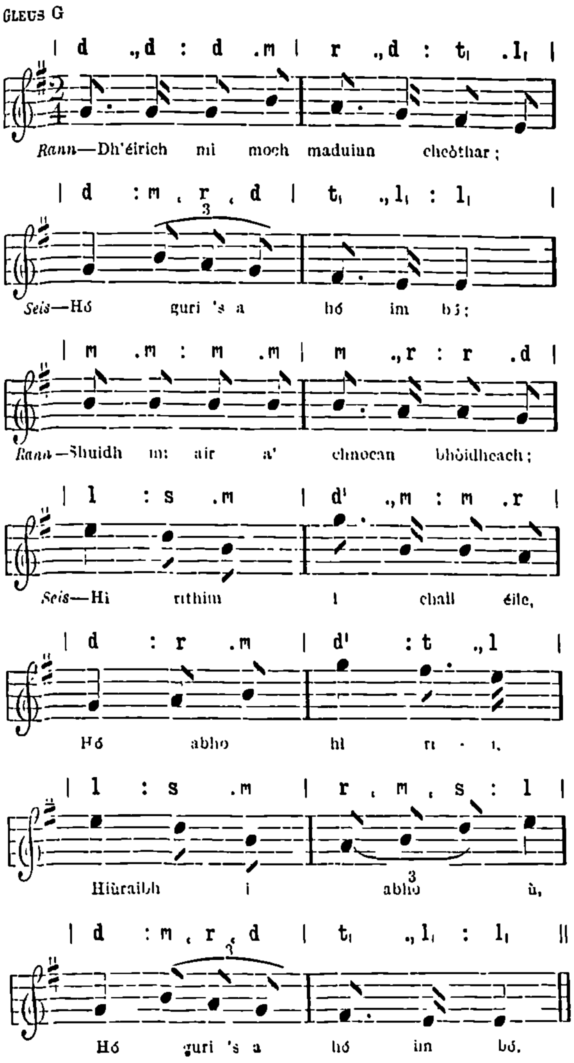
Tune 7


Cumha Chailein Ghlinn Iubhair.
Gleus B♭
Smaointean truagh a th'air m'aigne,
Dh'fhàg orm smuairean is airsneul:
An àm gluasad a'm leabaidh,
Cha chadal ach dùisg:
Tha mo ghruaidhean air seacadh,
Gun dion uair air mo rasgaibh,
Mu'n sgeul a chualas o'n Apuinn
Ghluais a' chaismeachd ud dunn.
Fear Ghlinn - Iubhair a dhìth ornn
Le puthar luchd - mioruin:
Mo sgeul dubhach r'a innseadh:
Thu bhi'd shineadh 'san ùir.
'S truagh gach duine de d' dhìlsean
O'n a chaidh do chorp prìseil
An ciste chumhainn, chaoil, dhaimaich,
'S ann an lion - anart ùr.

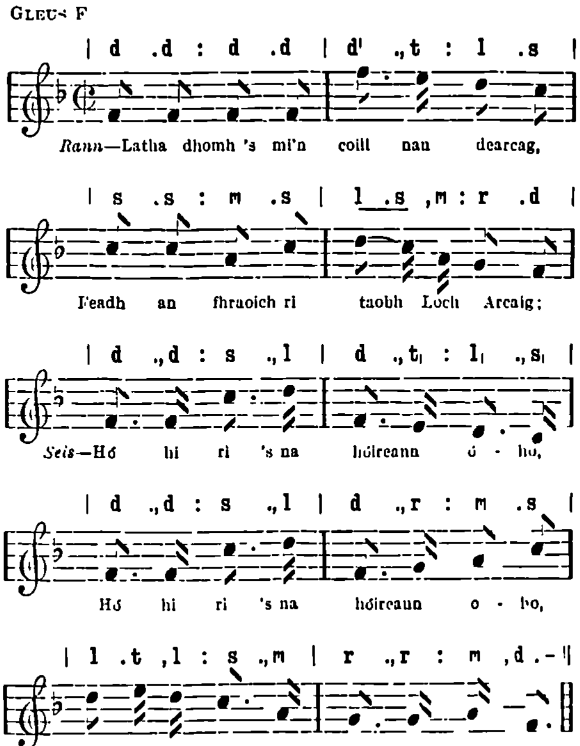
Tunes 8 & 9

(left and right, respectively)

Dh'éirich mi moch maduinn cheòthar.



Oran Chloinn Lachainn.

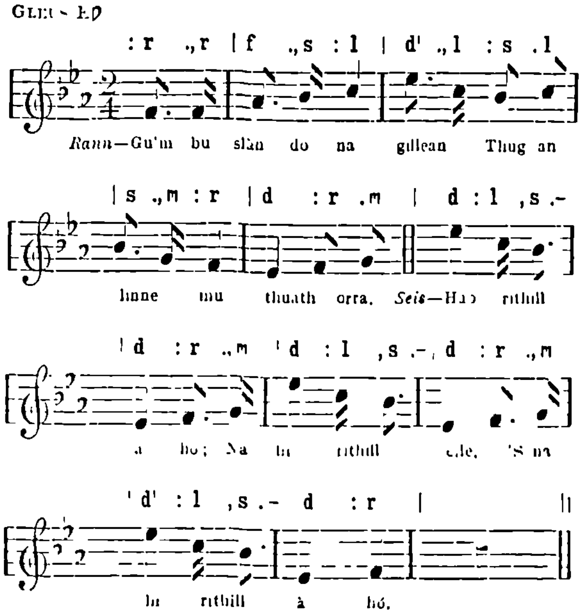


Tunes 10 & 11

(left and right, respectively)

Mo Roghainn 's mo Rùn.

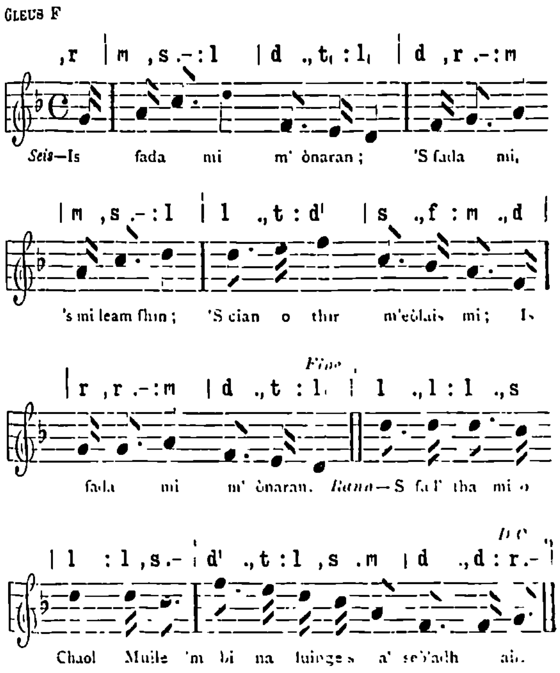
Gu'm bu Slàn do na Gillean.



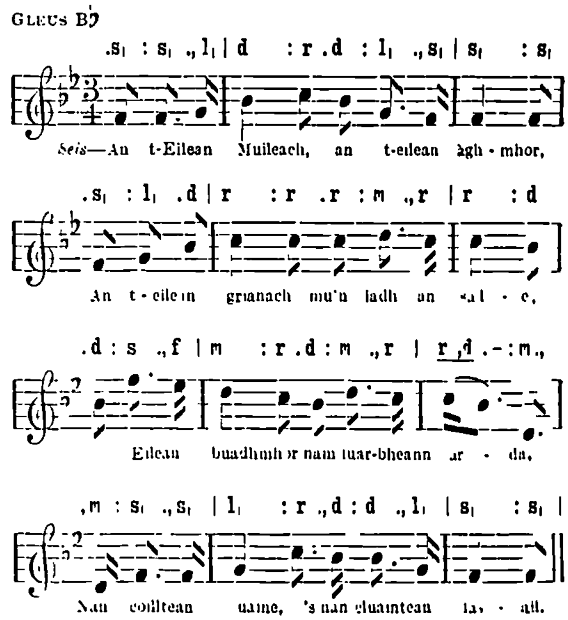
Tunes 12 & 13

(left and right, respectively)

Is Fada mi m' Ònaran.



An t-Eilean Muilcach.



Tunes 14 & 15

(left and right, respectively)

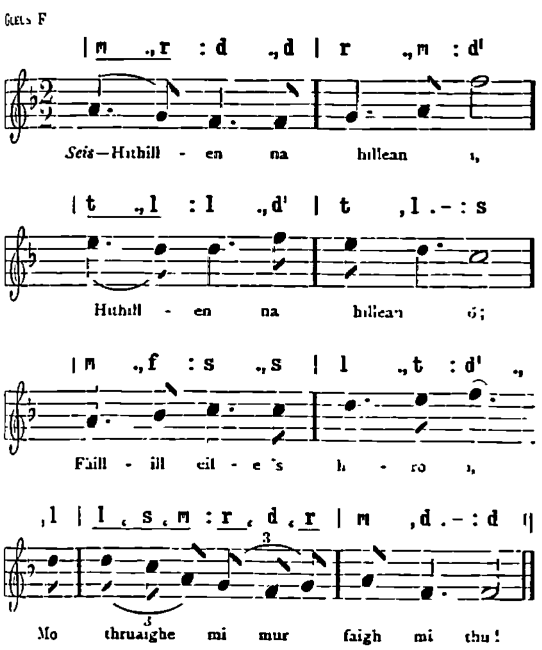
Màiri Bhòidheach.

A' Ghruagach Bhanail.

Tunes 16 & 17

(left and right, respectively)

'S e mo Cheist an Gille donn.

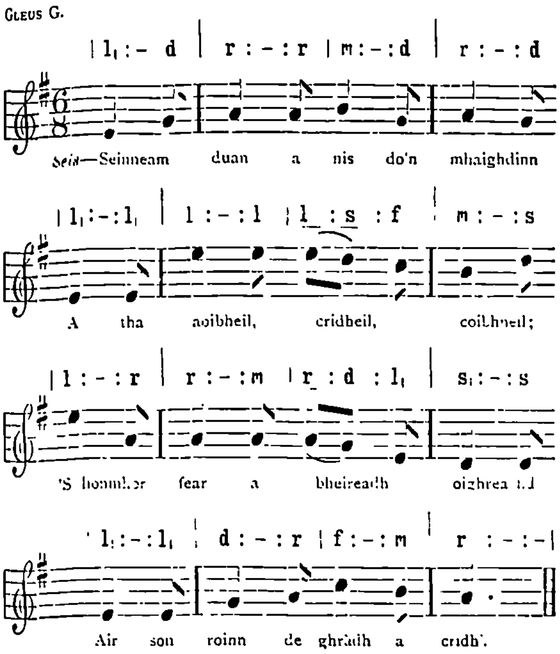


Soraidh.

Tunes 18 & 19

(left and right, respectively)

A' Mhaighdean Àluinn.



Eilidh Bhàn.

Tunes 20 & 21

(left and right, respectively)

Gur Trom, Trom mo Cheum.

Am Fonn.

Tunes 22 & 23

(left and right, respectively)

Latha Dhomh bhi 's Tigh-òsda.

Nigheanag a' chùil duinn, nach fhan thu?

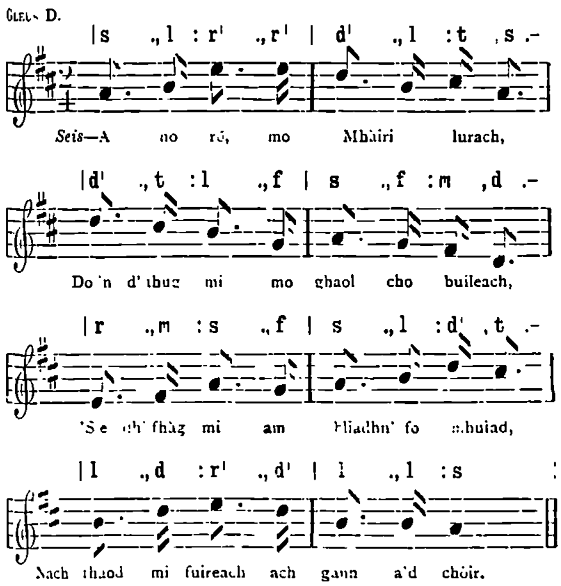
Tunes 24 & 25

(left and right, respectively)

Tunes 26 & 27

(left and right, respectively)

A ho ró, mo Mhàiri Lurach.



Mo Nighean Dubh.

Tunes 28 & 29

(left and right, respectively)

Oran do Iarla Bhraid-Albann.

Gleus B♭

Duanag an t-Seoladair.

Gleus F

Tunes 30 & 31

(left and right, respectively)

A' Bhanarach Dhonn a' Chruidh

GLEUS D

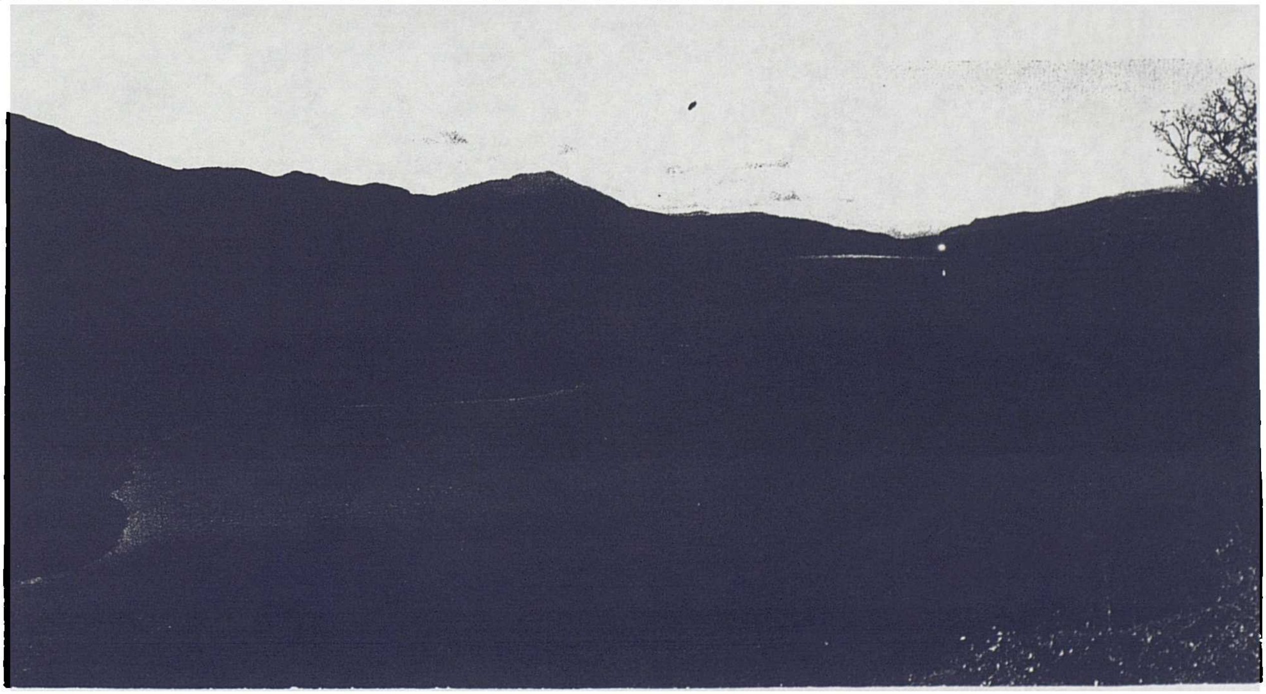
Fallan gun dìth thàinig thù.

GLEUS C

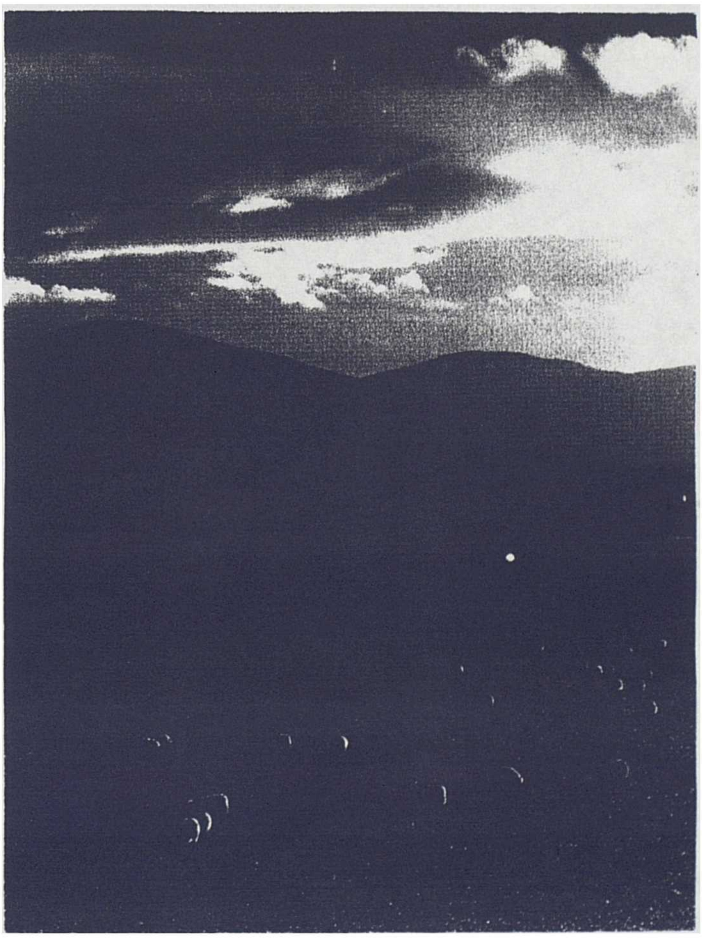
Appendix B: *14 Landscape picture pairs for Experiment Thirteen*

Pictures 1 & 2

(top and bottom, respectively)

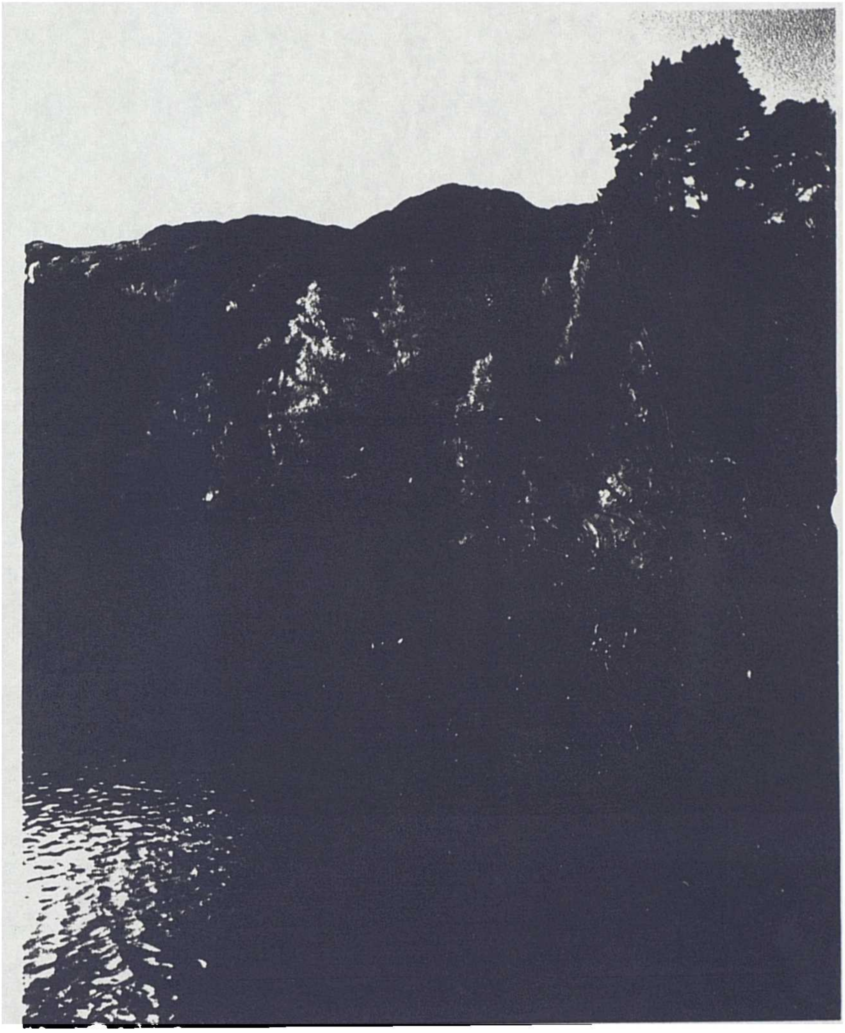


<u>Pictures 3 & 4</u>

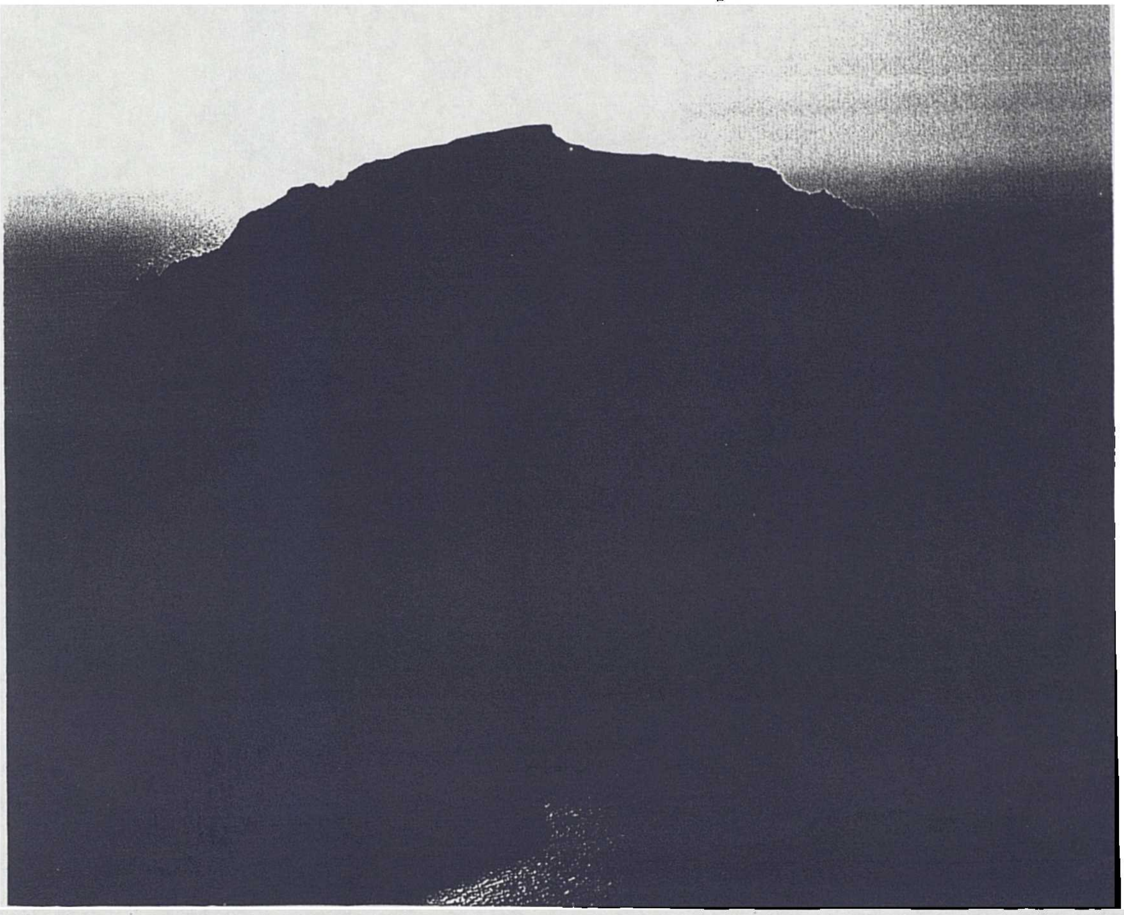


<u>Pictures 5 & 6</u>

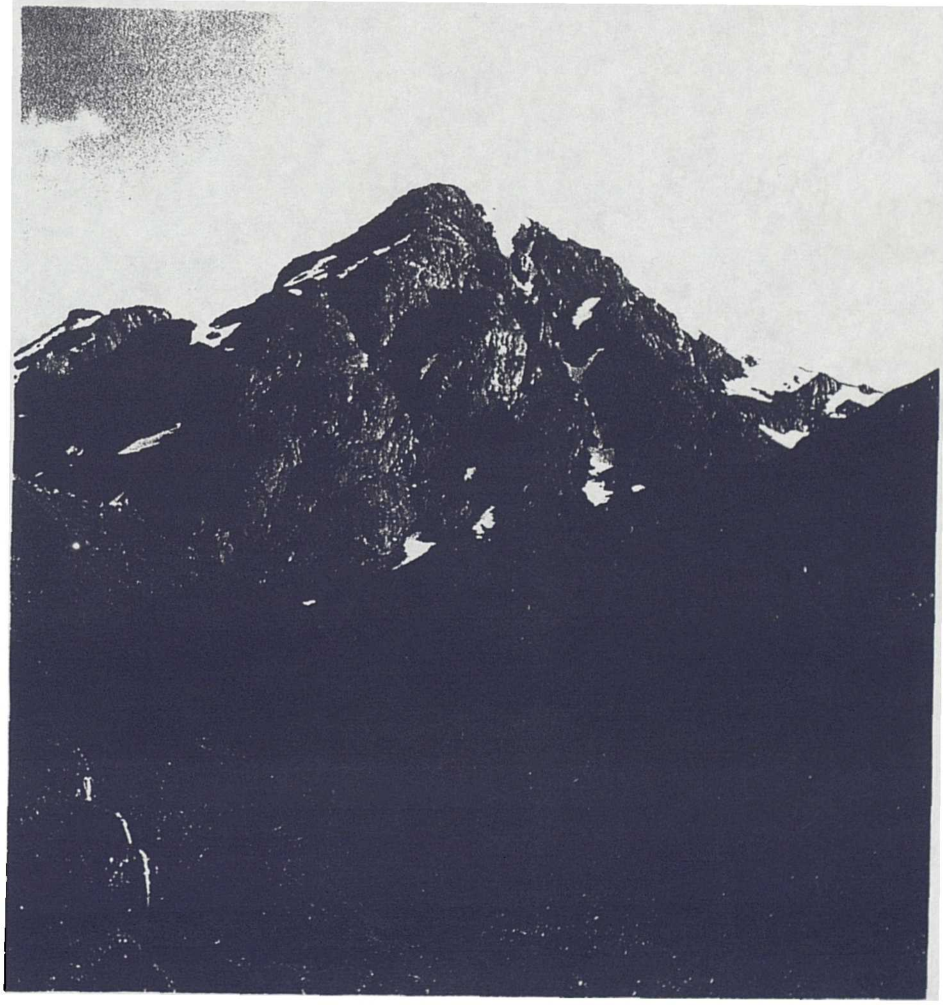
Pictures 7 & 8



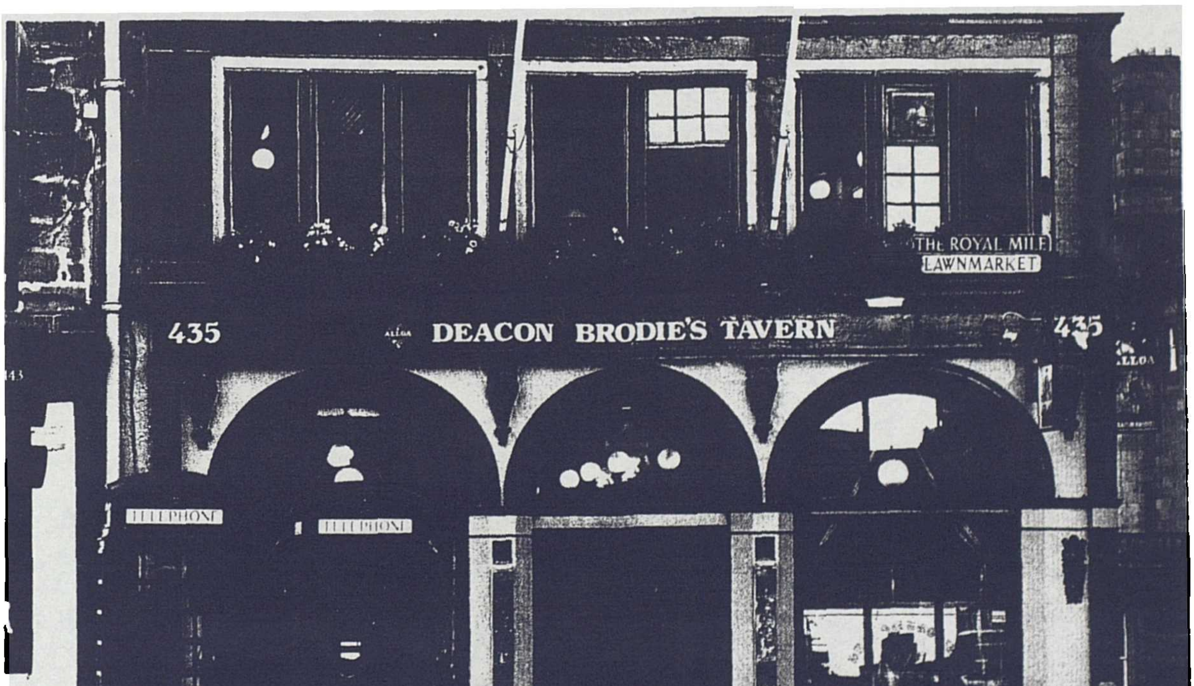
Pictures 9 & 10



Pictures 11 & 12

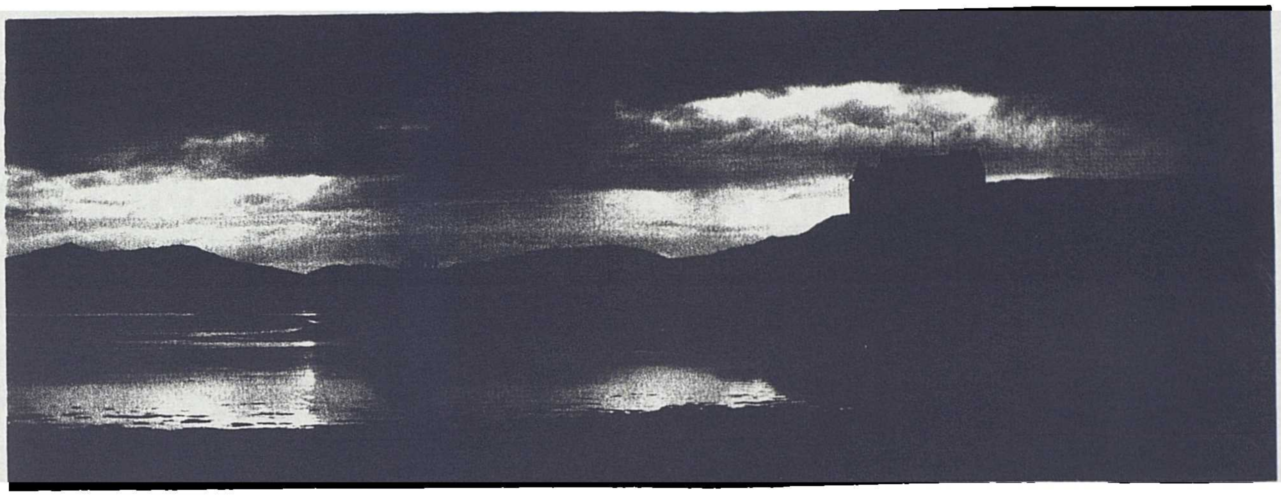


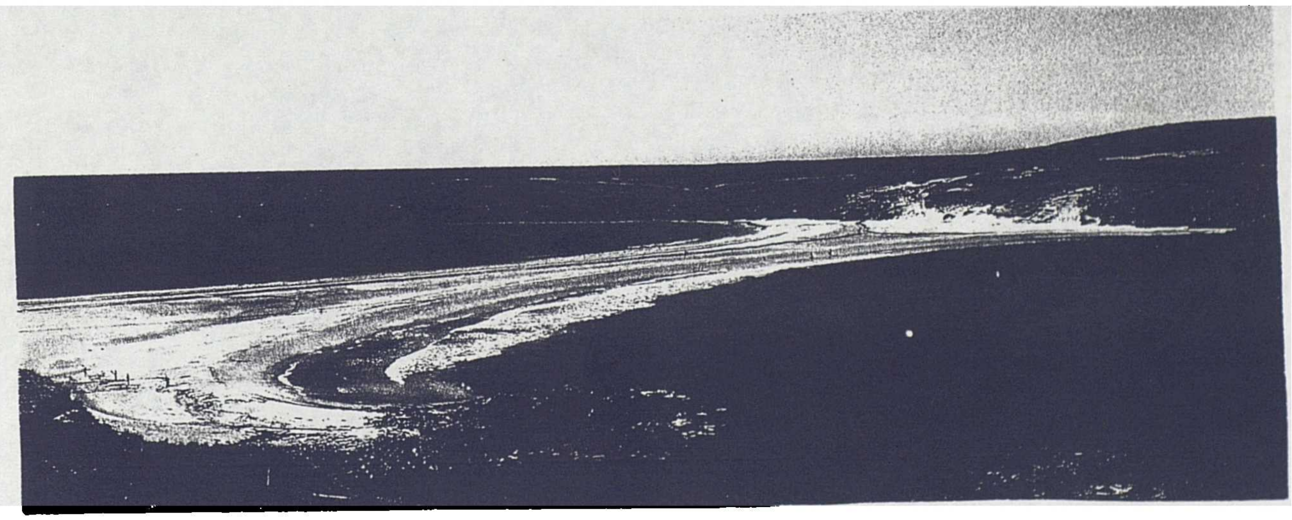
Pictures 13 & 14



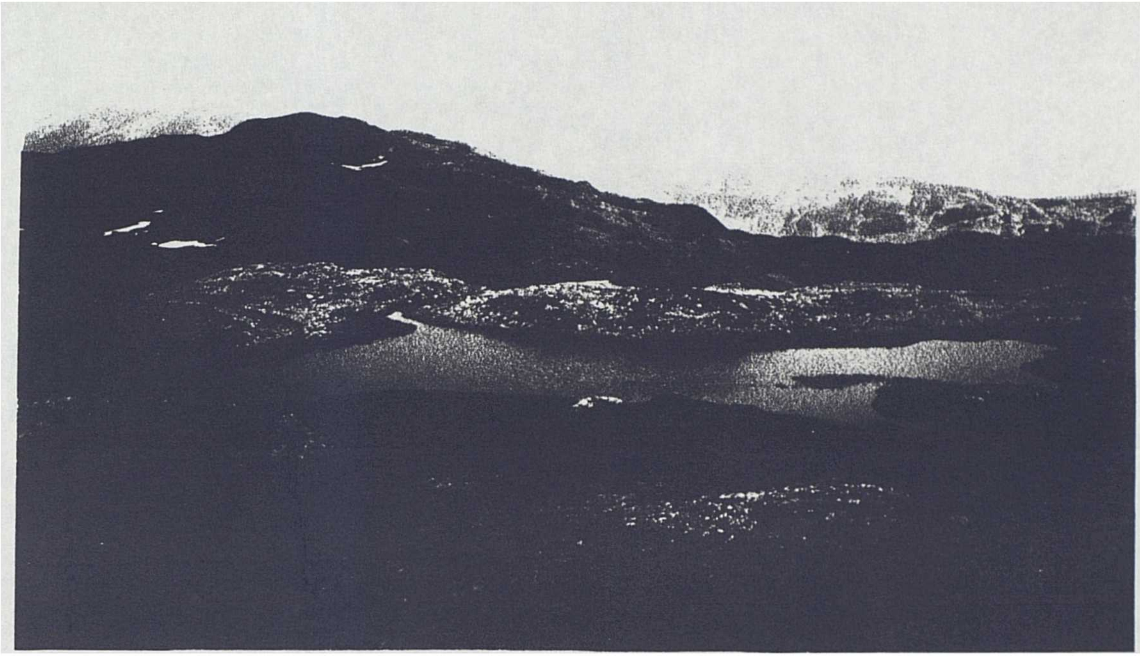
Pictures 15 & 16

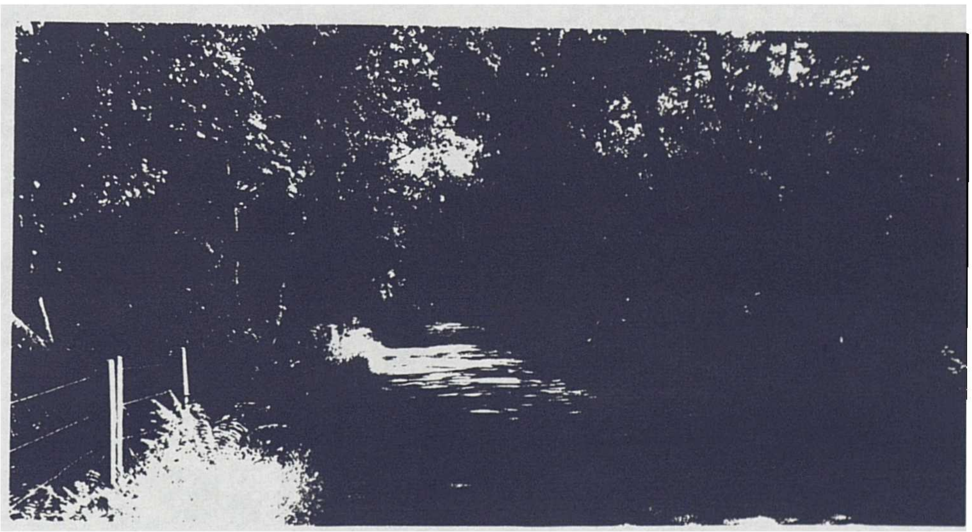
Pictures 17 & 18



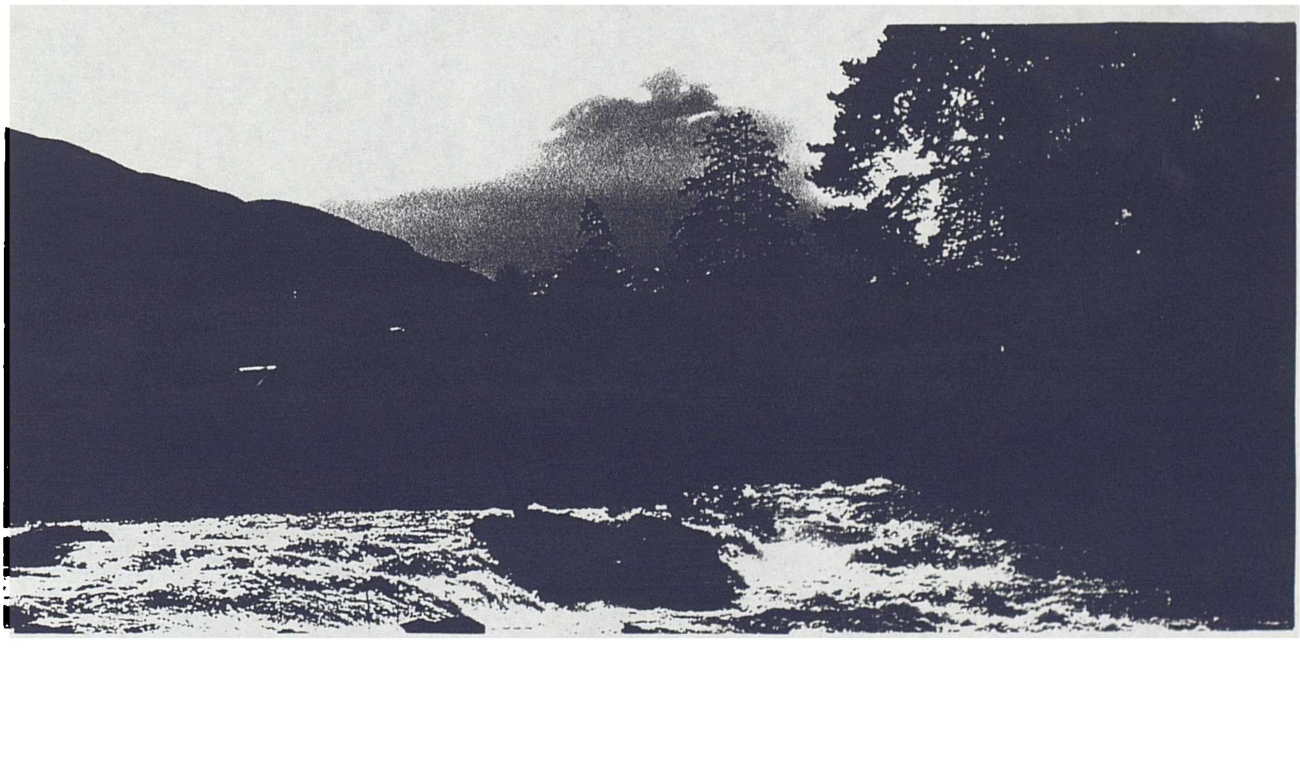


Pictures 19 & 20

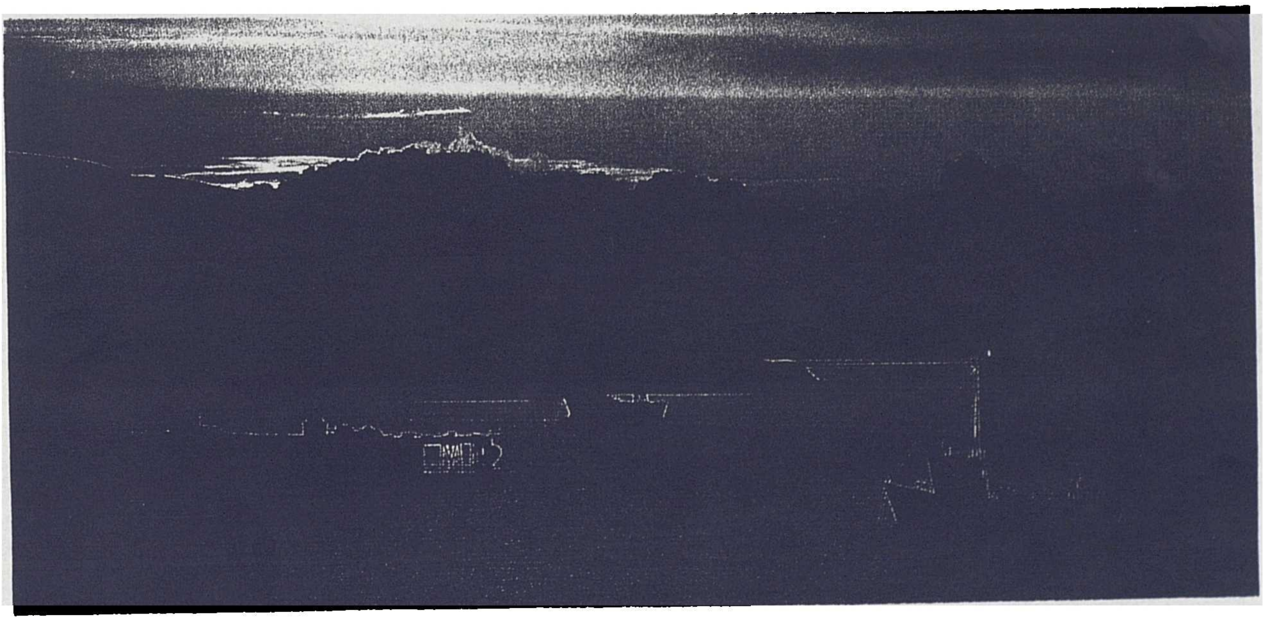




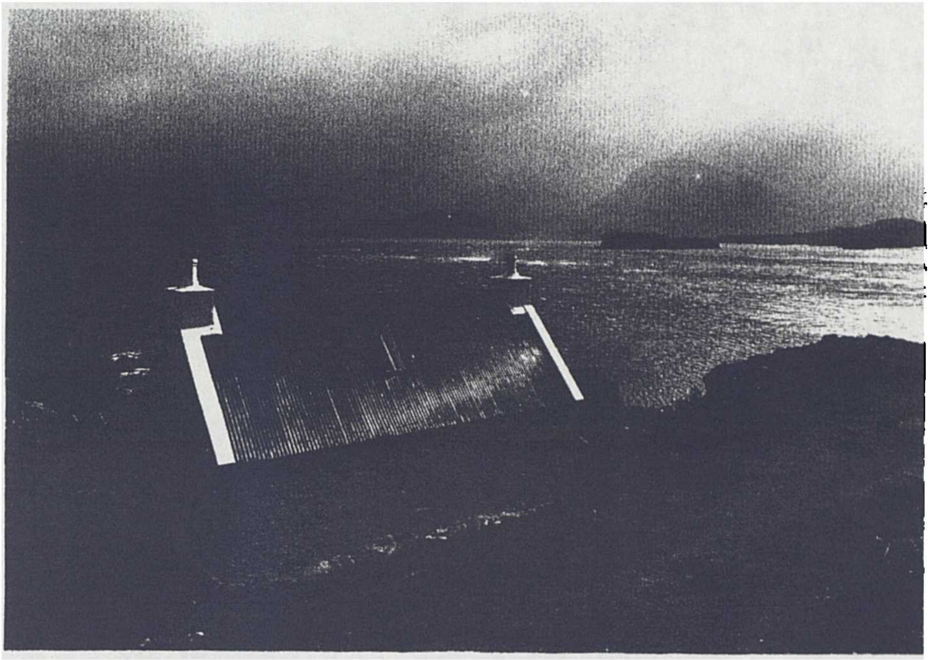
<u>Pictures 21 & 22</u>

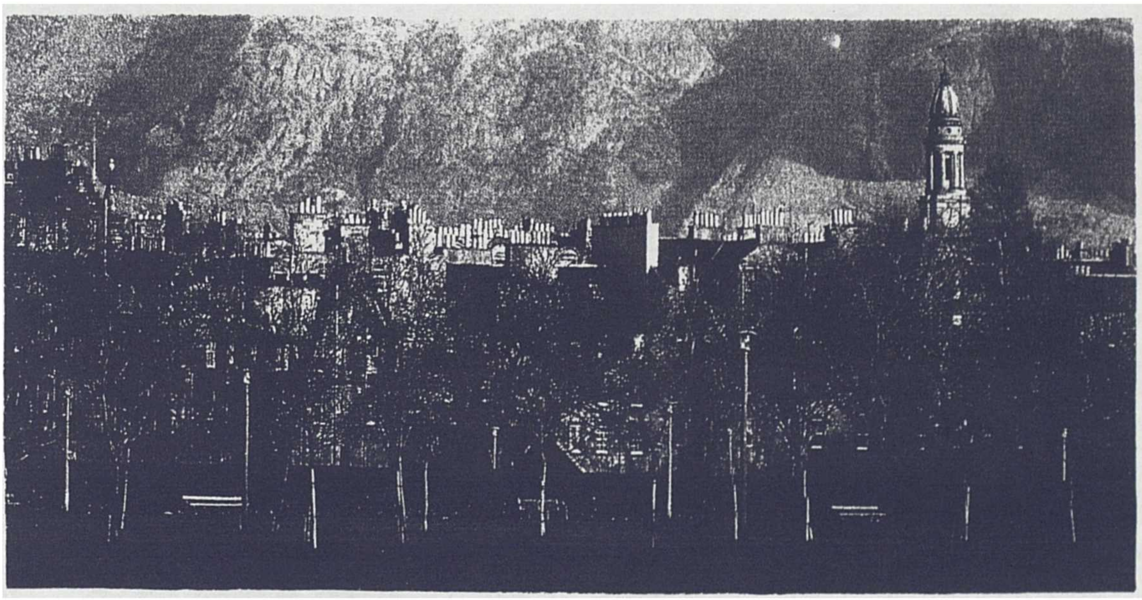


Pictures 23 & 24



<u>Pictures 25 & 26</u>





Pictures 27 & 28