How touch and hearing influence visual processing in sensory substitution, synaesthesia and cross-modal correspondences

Giles Hamilton-Fletcher

Thesis submitted for the degree of Doctor of Philosophy

University of Sussex

December, 2014

The thesis conforms to an ‘article format’ in which the middle chapters consist of discrete articles written in a style that is appropriate for publication in peer-reviewed journals in the field. The first and final chapters present synthetic overviews and discussions of the field and the research undertaken.

Chapter two is published in Multisensory Research as:

Hamilton-Fletcher, G., & Ward, W. (2013). Representing Colour in Sound and Touch for Sensory Substitution: A Review. *Multisensory Research*, 26, 503-532.

The author contributions are as follows: Giles Hamilton-Fletcher was responsible writing of the manuscript; Jamie Ward was responsible providing feedback and corrections to the manuscript.

Chapter three written in the style of an article appropriate for the Journal of Consciousness Studies

The author contributions are as follows: Giles Hamilton-Fletcher was responsible writing of the manuscript, experimental design, all aspects of data collection, data analysis; Jamie Ward was responsible providing feedback on the design and corrections to the manuscript.

Chapter four written in the style of an article appropriate for Multisensory Research

The author contributions are as follows: Giles Hamilton-Fletcher was responsible writing of the manuscript, experimental design, all aspects of data collection, data analysis and feedback on the experimental device; Thomas Wright was responsible for the initial concept of a touchpad based SSD named the Creole and the programming of the host and client applications for the experimental device in C# using Microsoft Visual Studio; Jamie Ward was responsible providing feedback on the experimental design, experimental device and corrections to the manuscript.

Chapter five written in the style of an article appropriate for Attention, Perception and Psychophysics

The author contributions are as follows: Giles Hamilton-Fletcher was responsible writing of the manuscript, experimental design, data collection for experiment 1 and 2, all data analysis; Christoph Witzel was responsible for the experimental program for experiment 1 and 2; David Reby was responsible for feedback on auditory stimuli; Devin Terhune and David Luke were also responsible for the study design for experiment 3; Mendel Kaelen and Robin Carhart-Harris were responsible for data collection for experiment 3; David Nutt was responsible for obtaining ethics for experiment 3; Jamie Ward was responsible providing feedback on the study and corrections to the manuscript.

Chapter six written in the style of an article appropriate for Multisensory Research

The author contributions are as follows: Giles Hamilton-Fletcher was responsible writing of the manuscript, experimental design, all aspects of data collection, data analysis as well as alterations to the Creole SSD to include CIE LUV / LCH colour space and output sound samples in through further programming in C# using Microsoft Visual Studio; Ward was responsible providing feedback and corrections to the manuscript.

Declaration

I declare that this thesis has not, in whole or in part, been submitted previously for the award of any degree at any University.

Signature:

Acknowledgements

Firstly I would like to thank my primary supervisor Jamie Ward who has provided excellent guidance in pursuing a wide variety of research topics while giving detailed feedback for me to improve myself. I am also grateful for the essential help in expanding my research beyond this thesis.

I would like to thank Anil Seth, Tom Wright and Keisuke Suzuki for their input on the technological side of the sensory substitution work. I am grateful to David Reby and Christoph Witzel for sharing their expertise on sound and colour research. Also thanks to Kristin Schmidt and Robyn Dillon for their assistance with testing and illustrations.

Thanks to all of the participants in my experiments for contributing their time and effort, and in particular the synaesthetes who provided such detailed descriptions of their experiences.

Importantly, thank you to D, for everything that made the lows bearable and the highs so much more enjoyable in our journey. I can only try to provide the same level of support to you, that you showed to me.

Finally I would like to dedicate this thesis to the memory of my mother, who helped provide the freedom and support necessary for me to follow this path.

Preface

Research reported here has also been presented at the following conferences: *Synaesthesia in Perspective: Development*, *Networks and Multisensory Processing Conference* (Hamilton-Fletcher and Ward, 2014); *British Academy Conference for Sensory Substitution and Augmentation* (Hamilton-Fletcher, Wright and Ward, 2013).

This research was funded by a three-year studentship from the ESRC and the University of Sussex

How touch and hearing influence visual processing in sensory substitution, synaesthesia and cross-modal correspondences

Giles Hamilton-Fletcher

Summary

Sensory substitution devices (SSDs) systematically turn visual dimensions into patterns of tactile or auditory stimulation. After training, a user of these devices learns to translate these audio or tactile sensations back into a mental visual picture. Most previous SSDs translate greyscale images using intuitive cross-sensory mappings to help users learn the devices. However more recent SSDs have started to incorporate additional colour dimensions such as saturation and hue.

Chapter two examines how previous SSDs have translated the complexities of colour into hearing or touch. The chapter explores if colour is useful for SSD users, how SSD and veridical colour perception differ and how optimal cross-sensory mappings might be considered.

After long-term training, some blind users of SSDs report visual sensations from tactile or auditory stimulation. A related phenomena is that of synaesthesia, a condition where stimulation of one modality (i.e. touch) produces an automatic, consistent and vivid sensation in another modality (i.e. vision). Tactile-visual synaesthesia is an extremely rare variant that can shed light on how the tactile-visual system is altered when touch can elicit visual sensations. Chapter three reports a series of investigations on the tactile discrimination abilities and phenomenology of tactile-vision synaesthetes, alongside questionnaire data from synaesthetes unavailable for testing.

Chapter four introduces a new SSD to test if the presentation of colour information in sensory substitution affects object and colour discrimination.

Chapter five presents experiments on intuitive auditory-colour mappings across a wide variety of sounds. These findings are used to predict the reported colour hallucinations resulting from LSD use while listening to these sounds.

Chapter six uses a new sensory substitution device designed to test the utility of these intuitive sound-colour links for visual processing. These findings are discussed with reference to how cross-sensory links, LSD and synaesthesia can inform optimal SSD design for visual processing.

Contents

1. Introduction

Listening to a vocal sound of a person speaking can influence much more than just auditory processing in the brain. This vocal sound can be integrated with the visual sight of seeing mouth movements to create a single unified event; one that in turn can influence the sound you thought you heard (McGurk & MacDonald, 1976). Different vocal sounds can produce specific perceptual colours in synaesthesia, a sensory condition where stimulation in one sense, automatically and consistently evokes sensations typically associated with another sense. For some synaesthetes, deep back vowels can bring to mind darker colours while front vowels may elicit more luminant colours (Marks, 1975). These synaesthetic tendencies can vividly illustrate connections between the senses that the wider population find intuitive and aesthetically appealing (Ward, Moore, Thompson-Lake, Salih & Beck, 2008), these implicit tendencies are collectively known as 'correspondences' (Spence, 2011). These tendencies can even be utilised to communicate colour information to the blind by using sound as its medium (Deville, Bologna, Vinckenbosch & Pun, 2009). This is seen in sensory substitution research, where dimensions normally associated with one sense (e.g. the horizontal and vertical position of luminance in vision) can be digitally recorded and then systematically converted to another sense (e.g. panning, pitch and loudness information in sound) for an end user to experience and learn to decode back into the original signal. All of these paths contribute to our understanding of multisensory processing, both in how the brain processes this information and how it can evaluate seemingly unrelated sensations against each other.

The first part of the introduction looks at how research into synaesthesia has managed to support personal claims of these vivid additional experiences. Methods of confirming the presence of synaesthesia are discussed, and from these 'confirmed synaesthetes,' their prevalence, heredity as well as associated neurological and behavioural characteristics. Different theoretical accounts of synaesthesia are discussed with relevance to different phenomenological types. This provides a basis to predict the perceptual effects of tactile-vision synaesthesia in Chapter three. The second part of the introduction looks into correspondence research and the wide variety of unconscious links discovered between a wide variety of senses. Correspondences have an important influence in the use of language through sound symbolism and also show intriguing similarities with synaesthesia. The potential mechanisms that can account for such a wide variety of sensory links are examined. Will these mechanisms be able to predict sound-colour correspondences investigated in Chapter five? The final part of the introduction looks at how sensory substitution research has used tactile and auditory pathways as alternative routes for visual information. The history, 'visual' achievements and limitations of widely used devices are discussed alongside their ability to recruit 'visual' regions for processing. These findings have implications for the brain as a flexible task-machine as well as what it means 'to see' philosophically. Chapter two looks at the expansion of colour within these devices, while Chapter four looks the importance of the presentation of this information while finally Chapter six examines how synaesthesia and correspondence research can lead to optimal sensory substitution design.

1.1. Synaesthesia

1.1.1. Defining synaesthesia and its variations

Developmental synaesthesia refers to a cognitive condition in which one type of stimulation (e.g. hearing sounds) induces an additional concurrent percept (e.g. a flash of colour) that is not normally experienced in the wider population (Grossenbacher & Lovelace, 2001; Simner, 2012). The concurrent experience does not replace the inducing stimulation, but is additional to it. Synaesthesia typically exists in individuals 'for as long as they can remember,' so experiencing colours for sounds appears as perfectly normal for such a person (Myers, 1911). There are many variations of synaesthesia and these are typically classified using inducer-concurrent pairings (e.g. sound-colour synaesthesia). These pairings can be seemingly intra-sensory such as grapheme-colour, where visually presented letters can elicit another visual percept. What appears to be intra-sensory at first glance however may also be revealed to also be elicited by higher-level representations of graphemes such as those delivered through hearing words or touching Braille (Simner, 2007; Stevens & Blakemore, 2004). Since grapheme and colour discrimination are two different modalities, this can also be described as cross-modal. Alternatively these pairings can be cross-sensory such as touch-colour, where stimulation in one sense elicits a percept in another sense (which would also be cross-modal as well). These inducer-concurrent links have several characteristics common to most synaesthetes: Firstly they are automatic, in that inducers almost always elicit concurrents for the synaesthete and that these concurrent experiences cannot be easily suppressed; Secondly they are typically consistent, in that specific inducer-concurrent pairings (e.g. 'a' elicits red) remain stable over the life of a synaesthete (Simner & Bain, 2013; also see Meier, Rothen & Walter, 2014). Specific pairings for groups of synaesthetes have general tendencies, such as more frequently used letters have a higher luminance, saturation as well as use more frequent colour terms (Beeli, Esslen & Jäncke, 2007; Simner et al., 2005). However these pairings are also highly idiosyncratic so that one synaesthete's coloured alphabet is quite different from the next. Synaesthesia is normally unidirectional such that sounds trigger colours but not vice versa, although a few bidirectional cases have been reported (Goller, Otten & Ward, 2009), and it may also be the case that bidirectional links occur implicitly by biasing behaviour but without eliciting a conscious concurrent experience (Rothen, Nyffeler, Von Wartburg, Müri & Meier, 2010).

There are several ways in which synaesthetes differ from other similarly classified synaesthetes. For example, the location of the visual concurrent can occur in multiple ways either internally in the mind's eye for 'see-associators,' a specific pairing can be known to be the correct for 'know-associators,' the visual concurrent can exist on the surface of the inducer for 'surface-projectors' and finally it can occur anywhere external to the body for 'near-space projectors' (Rothen, Tsakanikos, Meier & Ward, 2013; Ward, Li, Salih & Sagiv, 2007). Variations in inducers can also occur, such as with 'higher' and 'lower' synaesthetes (Dixon, Smilek, Cudahy & Merikle, 2000; Ramachandran & Hubbard, 2001). Higher synaesthetes have multiple forms of the same concept elicit the same colour, such as with TD, a synaesthete for whom '5' represented by graphemes, fingers or dice elicit the same cardboard-brown colour (Ward & Sagiv, 2007). Lower synaesthetes are more sensitive to the visual form so that variations in a graphemes contrast affect the subsequent photism (Hubbard, Manohar & Ramachandran, 2006). As suggested by this, synaesthetic photisms can also vary in their strength or vividness and this may be reflected in the level of cortical activity (Hubbard, Arman, Ramachandran & Boynton, 2005). The spatial-mapping can also vary between inducers and concurrents, such as with mirror-touch synaesthesia, where seeing another individual being touched creates an illusory sense of touch on their own body. This mapping can either be translated using a specular or anatomical frame of reference, so that viewing the face of another being touched on their left cheek, provides tactile sensations to the synaesthete's right or left cheek respectively (Banissy, Cohen Kadosh, Maus, Walsh & Ward, 2009). More controversially, synaesthesia may even vary in dimensions deemed to be 'core attributes' of synaesthesia such as consistency and automaticity. Exclusion from study based on failing to reach statistically significant levels of consistency or automaticity is typically done to avoid false-positives from non-synaesthetes managing to attain relatively high levels of consistency or automaticity. While individuals passing these tests can give researchers and sceptics confidence in an individual's synaesthetic condition, this may also exclude genuine synaesthetes or synaesthesias that do not pass this bar (Simner, 2012).

1.1.2. History of synaesthesia in science

Scientific interest in exploring synaesthesia is largely mirrored by the disciplines' focus on perception itself. The first reported case of synaesthesia is that of Georg Sachs in a medical dissertation on the subject of himself and his sister's albinism in 1812. In a chapter on the relationship between the eyes and colour, he describes his own synaesthesia through various instruments, notes, days and graphemes inevitably producing formless, coloured objects within his mind (Dann, 1998; Jewanski, Day & Ward, 2009). Eleven subsequent medical papers classified various forms of synaesthesia from 1849 to 1873, with multiple relatively large scale studies surveying the synaesthetic pairings of over seventy synaesthetes by Gustav Fechner (1871) and over seventy more by Bleuler and Lehmann (1881). Even on topics of individual synaesthesias such as sound-colour, Marks (1975) notes forty-four papers on the subject from this period until the 1940s. These studies sparked debate as to its neurological and learned origins, interestingly even initial explanations share some commonalities with contemporary theories (Jewanski, Simner, Day & Ward, 2011; Ward, 2013). However as synaesthesia describes a private subjective experience, this did not lend itself to behavioural testing. So with the rise of behaviourism and fall of consciousness studies in psychology, the subject of synaesthesia dipped in popularity until the 1980s (Marks, 1975; Van Campen, 1999). From a behaviourist viewpoint, synaesthesia could be considered little more than well learned associations between the senses (Howells, 1944). With the push to find objective tests of synaesthesia, consistency-testing of concurrents to inducing stimuli over weeks or months could be used to separate synaesthetes into a distinct category (Baron-Cohen, Wyke & Binnie, 1987). While even the extra-ordinary feats of consistency over time would not necessarily indicate altered perception, early neuroimaging research into synaesthesia found that experiences of colour concurrents shared neurological correlates with the perception of veridical colours (Aleman, Rutten, Sitskoorn, Dautzenberg & Ramsey, 2001; Nunn et al., 2002). This lends credibility to the notion that concurrent experiences can be perceptual-like since they utilise similar neural mechanisms, a point that cannot be accounted for by purely learned associations or mental imagery (Nunn et al., 2002). With increasing support for the reality of synaesthetic from neuroimaging studies (Rouw, Scholte & Colizoli, 2011), questions as to whether synaesthetic experiences were real, turned into how does synaesthesia operate? Contemporary questions look at how synaesthetic experience alters behaviour, perception, memory and neurology as well as how models of synaesthesia can account for the condition and inform models of non-synaesthetic perception (Ward, 2013).

1.1.3. Confirming synaesthesia

With questions regarding the reality of synaesthetic experience seemingly settled, consistency and automaticity testing is now used as a proxy for quickly and cheaply confirming the presence of synaesthesia. For grapheme-colour synaesthetes, colour consistency for individual graphemes is typically used, with the perceptual distance between a grapheme's colour at initial and subsequent compared in terms of colour difference. Examining distances between colours has become increasingly more nuanced and specific, with early comparisons using descriptive labels, and subsequently colour palettes and computer colour spaces such as RGB space (Asher, Aitken, Farooqi, Kurmani & Baron-Cohen, 2006; Baron-Cohen et al., 1987; Dixon, Smilek, Cudahy & Merikle, 2000; Eagleman, Kagan, Nelson, Sagaram & Sarma, 2007; Jordan, 1917). However, the use of colour models based on human colour perception continues to improve the sensitivity and specificity in discriminating synaesthetes (Rothen, Seth, Witzel & Ward, 2013). For automaticity testing, a synaesthetic Stroop interference task is commonly used (Mills, Boteler & Oliver, 1999). For grapheme-colour synaesthetes, an inducing letter is written either in its synaesthetic-colour or a different colour. When a letter's synaesthetic colour and written colour are congruent, there is a faster naming of both the photism and real colour relative to when these are incongruent (Dixon, Smilek & Merikle, 2004). Since improvements in consistency or automaticity are hard to fake for controls (although similar effects can be trained - Meier & Rothen, 2009), significant differences alongside reports of synaesthetic experience are taken as the gold standard in synaesthesia research.

1.1.4. Synaesthetic prevalence and genetics

With synaesthetes and non-synaesthetes discriminated using consistency measures, the prevalence and types of synaesthesia can be examined. Studies establishing synaesthesia using a random selection of the population found that 4.4% of the population have synaesthetic experiences, evenly split across the genders, with a little over 1% of the population experiencing grapheme-colour synaesthesia (Simner et al., 2006). This places synaesthesia as both more common and gender-neutral than previous studies which utilised self-reports, a method which suffers from self-report bias and having to make demographic assumptions (Baron-Cohen, Burtlf, Smith-Laittan, Harrison, & Bolton, 1996; Cytowic, 1997; Rich, Bradshaw & Mattingley, 2005; Ramachandran & Hubbard, 2001). For synaesthetes, approximately half of synaesthetes report experiencing more than one type of synaesthesia (Simner et al., 2006). Having one type of synaesthesia has a strong influence on what others may be experienced. For instance, letter-colour synaesthetes are more likely to experience colour from other sequences such as numbers, weekdays or months (Novich, Cheng & Eagleman, 2011). These produce several clusters of synaesthesias within individuals relating to coloured sequences, coloured music, coloured sensations, spatial sequences and non-visual sequelae which likely correspond to similar or close neural regions being affected by synaesthesia. Despite these tendencies, it is important to note one synaesthete can experience synaesthesias from multiple clusters and that relatives with synaesthesia may not have either the same types or clusters either. This indicates a common underlying mechanism for synaesthesia as a whole (Barnett et al., 2008a; Novich et al., 2011). Synaesthesia appears to have an inherited component with relatives of synaesthetes more likely to have synaesthesia than the general population (Barnett et al., 2008a). However while some specific genes are implicated in synaesthesia it is unclear what role they play at present (Asher et al., 2009; Tomson et al., 2011). Even in monozygotic twins one can have synaesthesia while the other does not, indicating an influence beyond genetics (Smilek et al., 2002; Smilek, Dixon & Merikle, 2005). As such any implicated genes may be quite broad in their effect on brain development such as through influencing cortical connectivity, with the specific phenotype determined through other influences (Bargary & Mitchell, 2008). One such potential environmental influence is through early associative learning between sensory stimuli, so experience of coloured alphabet sets could influence specific grapheme-colour links (Witthoft & Winawer, 2013; Yon & Press, 2014). Another possibility is through exposure to statistical regularities in the environment such as that smaller animals tend to produce higher pitched vocals, potentially influencing pitch-size links in both correspondences and synaesthesia (Mills, Boteler & Larcombe, 2003; Mondloch & Maurer, 2004). Overall it appears there is a heritable predisposition to synaesthesia with specific subtypes clustering together, but that genetics in of themselves do not play a purely determinist role, but rather interact with other influences.

1.1.5. Synaesthesia, perception and memory

Synaesthesia lends itself to a variety of behavioural changes, some of which can be accounted for by the presence of atypical cross-sensory links and others by a change in the sensory processes themselves. Inducing stimuli can benefit from the sensory processes of other modalities in synaesthesia. For instance, as temporal discrimination tasks are easier using auditory rather than visual stimulation (Guttman, Gilroy & Blake, 2005), the timing of visual flash sequences are easier to discriminate when accompanied by synaesthetic sounds for visual-auditory synaesthetes (Saenz & Koch, 2008). Likewise the addition of concurrents as an additional retrieval cue is one potential explanation of the increased memory abilities for word-lists seen in grapheme-colour synaesthetes, who show increased focus on the perceptual rather than semantic aspects of word-lists (Radvansky, Gibson & McNerney, 2011; Yaro & Ward, 2007). Memory advantages for grapheme-colour synaesthetes do not apply to all inducing stimuli such as with short-term memory tasks like digit-span (Rothen & Meier, 2010), but can be seen in non-inducing stimuli with enhanced visual memory for colour, rather than shape or spatial distribution (Pritchard, Rothen, Coolbear & Ward, 2013). This improvement through the concurrent modality is not restricted to memory either, with improved perceptual discrimination of concurrents, such as colour for grapheme-colour synaesthetes or touch for mirror-touch synaesthetes (Banissy et al., 2009). These concurrent advantages are also mirrored with structural changes to associated neural regions with larger grey matter densities and white matter connectivity (Banissy et al., 2012; Rouw et al., 2011). Functional changes also occur to the concurrent modality even without the use of inducers, with grapheme-colour synaesthetes showing an increased neural response in the parvocellular visual system and enhanced occipital excitability (Barnett et al., 2008b; Terhune, Tai, Cowey, Popescu & Cohen Kadosh, 2011). These changes in neural structure and functioning can also have negative effects on neighbouring neural regions, for instance a brain region implicated in visual motion perception, V5 / MT, for grapheme-colour synaesthetes show reductions in grey matter density alongside impairments in motion discrimination (Banissy et al., 2012; Banissy et al., 2013). As a result, the phenomenology of an individual's synaesthesia is tightly interwoven with changes in perception, memory as well as neural structure and functioning.

1.1.6. Theories of synaesthesia

There are multiple theories of how synaesthetic symptoms come about and how they relate to non-synaesthetic perception. These theories can largely be broken down into two sets. The first set of theories proposes that structural differences between the brains of synaesthetes and non-synaesthetes can alone produce synaesthesia. The second set instead suggests that functional differences create synaesthesia and that structural differences may be the result, rather than cause of, synaesthesia. Bargary and Mitchell (2008) further differentiate structural and functional theories into whether they use direct or indirect routes, creating four basic models of synaesthesia (see fig. 1.1).

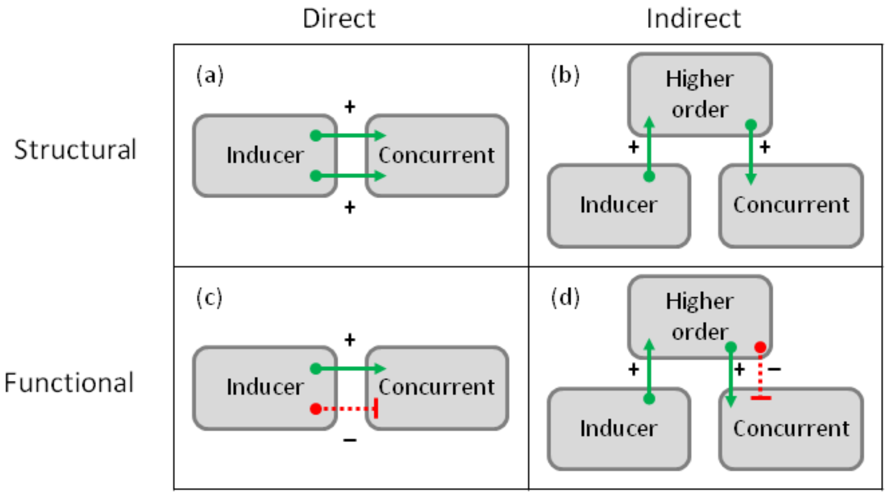


Figure 1.1. Models of synaesthesia. Models can either posit direct pathways between inducers and concurrents (a, c) or indirect through another area (b, d). The reason for the increased activation can either be through increased structural pathways (a, b) or through additional functional connectivity (c, d). Green arrows indicate excitatory connections (+), while red dashed arrows indicate functionally ineffective inhibitory connections (-). Figure adapted from Bargary and Mitchell, 2008.

The first structural theory is the cross-modal transfer hypothesis (Baron-Cohen, 1996; Meltzoff & Borton, 1979) which proposes that the regions utilised in synaesthesia are connected by atypical neural pathways, so that stimulation in the first modality inevitability stimulates the second. Maurer and Mondloch (2005) suggested the neonatal hypothesis, where any additional synaesthetic neural connections may be the result of atypical brain maturation. During infancy, activation in one modality can easily spread across modalities, creating a mixture of multiple senses for the child. As the brain matures, superfluous connections are 'pruned' and modalities are further differentiated, inhibiting this cross-activation. A lack of neural 'pruning' allows activity to cross the barrier between modalities and manifests itself as adult synaesthesia. Some theories are concerned with specific synaesthesias, such as the cross-activation theory's account of grapheme-colour synaesthesia (Ramachandran & Hubbard, 2001). They noted that the visual word form and colour discrimination areas neighboured each other, and suggested that a lack of neural pruning could result in grapheme sensitive neurons directly stimulating colour sensitive neurons. This theory yields support for projector synaesthetes with almost simultaneous activation in grapheme and colour sensitive regions and the influence of low-level grapheme-features on colour percepts (Brang, Hubbard, Coulson, Huang & Ramachandran, 2010; Brang, Rouw, Ramachandran & Coulson, 2011). However higher-level cognitions can still influence colours, with an identical graphemes representing different letters eliciting different colours (Cytowic & Eagleman, 2009, p. 75). Incorporating these influences leads to theories such as the cascaded cross-tuning model, where low-level features elicit colour that is subsequently fine-tuned with higher-level processing of the grapheme (Brang et al., 2010). Structural accounts of synaesthesia also have supporting evidence from increased neural pathways between relevant brain regions in synaesthetes (Rouw & Scholte, 2007). Increased structural connectivity between the thalamus and somatosensory regions has resulted in acquired sound-touch synaesthesia (Ro et al., 2007). Beyond the additional connectivity, there are structural and functional differences to regions affected by synaesthesia (Barnett et al., 2008b; Rouw & Scholte, 2010; Rouw et al., 2013). Contrary to the necessity of altered structure however is the presence of synaesthesia-like phenomenology in drug-induced or hypnotic contexts for non-synaesthetes (Cohen Kadosh, Henik, Catena, Walsh & Fuentes, 2009; Sinke et al., 2012). However certain behavioural consistencies with synaesthesia do not appear to be induced through hypnosis (Anderson, Seth, Dienes & Ward, 2014).

Theories that propose functional changes as the cause of synaesthesia include the disinhibited feedback model (Grossenbacher & Lovelace, 2001), where higher-level multisensory regions between the inducing and concurrent modalities produce disinhibited feedback to the concurrent regions, raising the concurrent modalities neural excitation. One potential source of disinhibition is via serotonergic pathways (Brang & Ramachandran, 2008), interestingly, serotonin agonists are also implicated in similar experiences through hallucinogenic drug-induced synaesthesia (Luke & Terhune, 2013). Another functional explanation is that of the hyperbinding model, where normal binding mechanisms in the parietal lobe become hyperactive in synaesthetes binding the inducer and concurrent together, supporting this is the reduction of interference in synaesthetic Stroop tasks during parietal inhibition through TMS (Esterman, Verstynen, Ivry & Robertson, 2006; Muggleton, Tsakanikos, Walsh & Ward, 2007). Anomolies of cross-sensory binding can produce the double-flash illusion, where two auditory beeps and a single visual flash can produce a second illusory flash. While some evidence has suggested that grapheme-colour synaesthetes experience more of these illusions indicating a general hyper-sensitivity to binding (Brang, Williams & Ramachandran, 2012), other research has found that groups of grapheme-colour and sound-colour synaesthetes are less likely to experience the double-flash illusion and have a smaller temporal window when they do, which instead suggests a reduced binding relative to controls to produce more accurate perceptions (Neufeld, Sinke, Zedler, Emrich & Szycik, 2012). This suggests the possibility for specific types of synaesthesia to have different influences on cross-sensory binding. Some theories focus on information processing such as the re-entrant hypothesis for grapheme-colour synaesthesia (Smilek, Dixon, Cudahy & Merikle, 2001). Which proposes that as visual stimulation is broken down into components such as form and colour for processing, in projector synaesthesia there is additional feedback from the posterior inferotemporal cortex / fusiform areas into V4 influencing visual perception. This could help explain how a synaesthetic photism could 'mask' the grapheme inducing it on congruently coloured backgrounds in visual search tasks.

Different theories may be better equipped to explain some synaesthetic types over others. The prevalence of different types of synaesthesia is also indicative of the likelihood of certain theories. The higher prevalence of synaesthesias in which the inducing and concurrent modalities are close together with additional neural connections indicate that local connectivity theories for synaesthesia may be a common pathway (Day, 2014; Hubbard, Brang & Ramachandran, 2011; Rouw & Scholte, 2007). For synaesthesias such as sound-colour where the modalities are far apart, frontal connectivity and parietal binding mechanisms appear to be more involved than direct cross-activations between modalities (Neufeld et al., 2012; Zamm, Schlaug, Eagleman & Laui, 2013). Finally, the spatial frame in which the concurrents are involved in is also important. Internal representations in associator synaesthetes show a parietal top-down mechanism of functional connectivity, best accounted for by disinhibited feedback models, while external representations in projectors show bottom-up fusiform activations in concordance with cross-activation theories (Van

Leeuwen, den Ouden & Hagoort, 2011). As such, different theories of synaesthesia may not be in conflict, but rather, best suited to specific types.

1.1.7. Synaesthesia and the present research

Whether the recruitment of additional processes can be deemed the result of functional or structural changes, the extent to which concurrent processes can alter the perception of the inducer is unknown. Saenz and Koch (2008) found that motion-hearing synaesthetes are superior at temporal discrimination for flashing lights since the additional auditory concurrent can utilise the superior temporal discrimination abilities of the auditory cortex. However, this does not tell us whether they have superior *visual* temporal discrimination. By contrast, the tactile modality is already known to recruit visual-orientation processes in non-synaesthetes during tactile-orientation acuity tests (Zhang et al., 2005). In addition, visual deprivation allows the recruitment of the visual cortex for tactile-distance discrimination tasks (Merabet et al., 2008). Tactile-vision synaesthesia allows us to investigate the effect of synaesthetic visual recruitment on these processing routes. Will synaesthetes show a strengthening of existing tactile-visual connections we all have or utilise links normally masked by normal visual stimulation? This also allows an opportunity to examine the phenomenology of one of the rarer variants of synaesthesia to gauge the tendencies and predict the mechanisms used to translate tactile into visual stimulation (Day, 2014). These findings also have implications for the processing of tactile stimulation after visual cortex recruitment resulting from tactile-visual sensory substitution expertise or visual phenomenology from touch (Kupers et al., 2006; Ortiz et al., 2011; Ward & Wright, 2014). Previous investigations into tactile-vision synaesthesia have also shown that changes in inducing stimulation can alter the concurrent phenomenological experience in ways similar to tactile-visual correspondences in the wider population (Ludwig & Simner, 2013; Simner & Ludwig, 2012). Chapter three also examines whether a previously observed weight-luminance correspondence is also mirrored in synaesthetic photisms, potentially indicating shared mechanisms (Walker, Francis & Walker, 2010). Both correspondence research and its relationship to synaesthesia are explored further in the next section.

1.2. Cross-modal correspondences

1.2.1. Definition and behavioural effects

If forced to answer the question "is a high pitched note more similar to white or to black?" most individuals will intuitively match high pitch with a higher luminance, like white (Martino & Marks, 1999; Ward et al., 2006). Correspondences include preferences for matching low-level stimulus features in one modality (e.g. High-pitch) to features in another modality (e.g. High-luminance). These cross-modal matching tendencies are widespread across the population and relative, in that they are influenced by context of the stimulus. For example, an 800Hz tone would be 'high' in a 100-800Hz range, but 'low' in an 800-3200Hz range. This is different to synaesthesia, which displays more absolute mappings, with specific tones consistently related to specific colours irrespective of context (Thornley Head, 2006; Ward et al., 2006). Some correspondences are matched between sensory dimensions, but the direction of the correlation depends on the individual. For instance, loudness has been mapped to either increased or decreased brightness in specific individuals (Marks, 1974). With these correspondences, congruent pairings are processed faster relative to incongruent pairings (Martino & Marks, 1999). These tendencies are largely unconscious as well as being bi-directional so that high-luminance also helps in predicting and processing high pitch (Evans & Treisman, 2010). Many of these correspondences are present in infancy, such as with pitch-height and pitch-sharpness correspondences (Walker et al., 2010), or in close evolutionary relatives, with pitch-luminance correspondences shown through their accuracy in cross-modal tasks in chimpanzees (Ludwig, Adachi & Matsuzawa, 2011). However since there is no difference in reaction times for congruent / incongruent matching for the Chimpanzees, these correspondences may not operate in fundamentally the same fashion as similar correspondences in humans that do show reaction time differences. As such, some correspondences appear to occur either innately or at least independent of language development. As mappings between auditory and visual characteristics are shared between individuals, correspondences have also been investigated as the scaffolding to support utilising sound symbolism to represent objects through language (Ramachandran & Hubbard, 2001). Higher-level correspondences can also exist, where matching is done with reference to an intermediary process, such as through shared semantic meanings, words, concepts or emotions (Palmer, Schloss, Xu & Prado-León, 2013; Spence, 2011). As many of these are based on language, experience or cultural associations, learning can play a strong role in shaping these.

Both high and low-level correspondences can influence a variety of cognitive processes with differing effects. One such process is audio-visual integration where two incoming forms of information are bound together as a single object or occurrence. Most research into audio-visual integration has focused on the extent to which congruency in temporal and spatial factors lead to unified events (Calvert, Spence & Stein, 2004; Spence & Driver, 2004). However correspondences can also influence binding. This can be seen using temporal order judgement tasks, where increased binding is reflected through increased difficulty in judging which sensory stimulation occurred first. For high-level correspondences, multisensory binding can occur when visual and vocal gender cues are matched, which makes the identification of whether visual or auditory stimulation came first more difficult (Vatakis & Spence, 2007). Likewise, lower-level correspondences with congruent pairings such as higher pitched tones with smaller sizes also affect the temporal and spatial binding window (Parise & Spence, 2008, 2009). Low-level matches seem to affect processing in later stages than the vocal ventriloquism effect (Keetels & Vroomen, 2011). As such, higher and lower-level correspondences may affect multisensory integration differently in unifying multisensory signals (Vatakis, Ghazanfar & Spence, 2008). Low-level correspondences can also influence visual attention to higher or lower locations using high and low pitch respectively, an effect which can be influenced through cognitive control suggesting this influence occurs after sensory encoding at a later processing stage (Chiou & Rich, 2012). Alternating pitches from high to low or the reverse do not appear to affect sensory encoding during vertical motion perception tasks unless there is prior pitch-space training (Hidaka, Teramoto, Keetels & Vroomen, 2013). Likewise, pitch-luminance associations only aid visual searches for congruent targets when top-down cognitive control is facilitated through explicit explanations of this correspondence (Klapetek, Ngo & Spence, 2012). However even arbitrary associations learnt through repetition can show similar perceptual processing effects to low-level correspondences or synaesthesia (Colizoli, Murre & Rouw, 2012; Elias, Saucier, Hardie & Sarty, 2003; Meier & Rothen, 2009; Rothen, Wantz & Meier, 2011). As such, both innate and learned correspondences appear to affect perceptual processing but only training and high-level correspondences appear to influence perceptual judgements. From this, correspondences learned through statistical regularities in the environment could also provide an evolutionary advantage in extracting unified coherent objects from multi-sensory stimulation.

1.2.2. History of correspondences - symbolism, synaesthesia and other senses

Early work on cross-modal correspondences was primarily focused on sound-symbolism, where it was noted that visual characteristics of objects were reliably matched to specific nonsense words. Köhler (1929, 1947) found that rounded objects were associated with "baluma" or "maluma," while pointed objects were associated with "takete." Sapir (1929) likewise reported that when asked to match specific vowel sounds with shapes, over 80% of participants matched the /i/ in "mil" consistently to smaller circles while /a/ in "mal" was matched to larger ones. Sapir proposed that either the acoustic properties or vocal movements required to make the sounds were the basis of their visual correspondences. Newman (1933) expanded this association between vowels and verbal descriptors of size and colour, finding that back vowels related to larger sizes and darker luminances, and that this association remains stable across a 9 to 16 age range. This early research on associations between vowels and visual features was primarily referenced with note to language; however parallel research on vowel-size and vowel-colour correlations were also occurring with early synaesthesia research (Collins, 1929; Coriat, 1913; Ginsberg, 1923; Reichard, Jakobson & Werth, 1949; Riggs & Karwoski, 1934; Rose, 1909; Marks, 1975; Myers, 1911, 1914). Similar directions in the links for both correspondences and synaesthetes in vowel-colour, vowel-size and pitch-luminance correspondences (Miyahara, Koda, Sekiguchi & Amemiya, 2012; Moos, Simmons, Simner & Smith, 2013; Rojczyk, 2011; Ward et al., 2006; Wrembel, 2009) underscored similarities between synaesthetes and the wider population in their audio-visual mappings (see fig. 1.2). Playing up the similarities between synaesthesia and correspondence research can be seen from the monikers used to describe correspondences as synaesthetic correspondences or associations or perhaps most explicitly, as weak synaesthesia (Braaten, 1993; Martino & Marks, 2001; Melara & O'Brien, 1987; Parise & Spence, 2008; Walker et al., 2010).

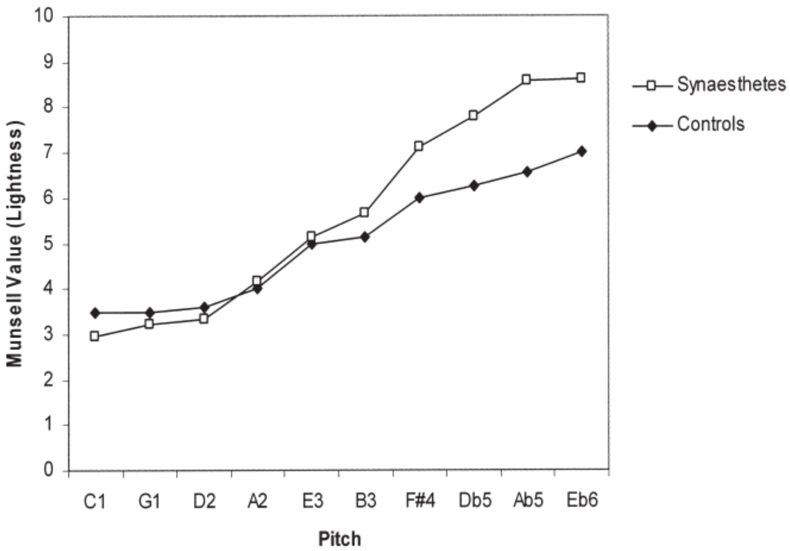


Figure 1.2. Relationship between pitch and luminance for both synaesthetic photisms and correspondences chosen by non-synaesthetes in free choice selections. Higher Munsell values indicate more luminant colour selections. Both synaesthetes and non-synaesthetes choose similar directions in their associations between pitch and luminance, indicating common underlying mechanisms (figure from Ward et al. 2006).

Examining the relationship between correspondences and synaesthesia has been a contentious issue, particularly with regards to whether both exist on a continuum or whether they are discrete entities despite some surface level similarities. Martino and Marks (2001) originally coined the descriptive term 'weak synaesthesia' to describe correspondences, suggesting that both express the same tendencies in differing ways, with 'strong synaesthesia' providing vivid imagery while 'weak synaesthesia' influences perceptual processing, similarity judgements and language. Deroy and Spence (2013) suggest that this positioning is misleading as it presumes a priori that they differ on one continuous dimension despite extensive differences between them (see table 1.1). They argue that multiple differences are binary (e.g. uni/bidirectionality) and so cannot vary along this continuum. A continuum also suggests a middle ground, such as with inconsistent conscious experiences to inducers, an aspect that currently lacks evidence (Deroy & Spence, 2013a). Correspondences might also serve a different purpose to synaesthesia, internalising rules from the environment using a system that remains malleable, aiding in multisensory perception and integration, whereas synaesthesia, once present, is less sensitive to environmental statistics. One potential danger of the continuum model is that it may prematurely blur the lines between correspondence and synaesthesia research. A vivid illustration of this tendency is to classify the unconscious priming of synaesthetic inducers from concurrents as 'implicit bidirectionality' rather than as correspondences that reflect the statistics of that synaesthete's unique internal experience (Deroy & Spence, 2013a; Gebuis, Nijboer & Van der Smagt, 2009; Rothen, Nyffeler, Von Wartburg, Müri & Meier, 2010). Correspondences, synaesthesia and 'implicit bidirectionality' all appear to use similar parietal regions, so the idea of shared underlying mechanisms does have some support, however they cannot account for the conscious experiences (Bien, ten Oever, Goebel & Sack, 2012; Esterman, Verstynen, Ivry & Robertson, 2006; Gebuis et al., 2009; Muggleton, Tsakanikos, Walsh & Ward, 2007; Rothen et al., 2010). Taken alongside evidence of similar matching tendencies between the senses, synaesthesia is likely to use a combination of correspondence-mechanisms and early experience to inform their phenomenology (Simner et al., 2005; Ward et al., 2006; Witthoft & Winawer, 2013).

Table 1.1. Similarities and differences between synaesthesia and cross-modal correspondences (table adapted from Deroy & Spence, 2013a)

	Synaesthesia	**Correspondences**
Similarities	Cross-modal	Cross-modal
	Surprising	Surprising
	Consistent	Consistent
Differences	Rare	Frequent
	Mostly Idiosyncratic	Universal
	Automatic	Some automatic, some not
	Conscious	Unconscious
	Absolute mappings	Relative mappings
	Unidirectional	Bidirectional
	Intransitive	Transitive
	Rigid	Malleable
	Not trainable	Trainable
	No animal evidence yet	Animal evidence
	Little infant evidence	Infant evidence

Due to the universal attribute of correspondences, they have the potential to contribute to effective and intuitive verbal communication. Building on the work of Köhler (1929) and Sapir (1929) on audio-visual correspondences, Ramachandran and Hubbard (2001) suggested that the 'bouba-kiki effect,' where over 95% of participants associate the sound of "bouba" with round shapes and "kiki" with angular shapes, had important implications for sound symbolism in the evolution of language. They suggested that these shared rules helped in developing proto-languages as well as serving as the basis for metaphor. The 'bouba-kiki' effect is also seen in remote cultures, suggesting that this correspondence does not require specific languages (Bremner et al., 2013). Even infants make sound-space and sound-shape correspondences to full words and with increasing familiarity with language, older individuals transfer these associations to individual vowels and consonants, allowing the ability to elicit correspondences with new words (Ozturk, Krehm & Vouloumanos, 2013). Westbury (2005) showed that these sound-shape associations appear to operate at a pre-semantic level with congruent pairings speeding up perceptual processing. There are only a few effects known to disrupt these correspondences, first is the presence and severity of autism (Oberman & Ramachandran, 2008; Occelli, Esposito, Venuti, Arduino & Zampini, 2013), second is through damage to the left angular gyrus within temporal-parietal-occipital regions (Ramachandran & Hubbard, 2001, 2003; Ramachandran, Azoulai, Stone, Srinivasan & Bijoy, 2005). This provides initial evidence that the correspondences used in sound-symbolism may operate in higher-level multisensory regions.

While audio-visual correspondences have important implications for language, correspondences can refer to any cross-sensory links. The wide variety of cross-sensory links observed help to reveal common processes, patterns and effects that underlie cross-sensory matching. Aspects of touch have been linked to vision with luminance related to vibration, pressure, roughness, softness and roundness of tactile sensations, while saturation is related to smoothness and softness, finally hue has been related to softness and temperature (Ludwig & Simner, 2013; Martino & Marks, 2000; Morgan, Goodson & Jones, 1975; Ward, Banissy & Jonas, 2008). The sense of smell and taste have also received a huge focus, being "notoriously difficult to describe," (Crisinel & Spence, 2011, p. 151) conveying the sense of smell or taste in other sensory terms becomes a useful tool in priming and describing these sensations (Crisinel et al., 2012; Spence, Puccinelli, Grewal & Roggeveen, 2014). Smell has been related to auditory pitch (Belkin, Martin, Kemp, & Gilbert, 1997; Piesse, 1891; Von Hornbostel, 1931), complex sounds (Seo & Hummel, 2011), instruments (Crisinel & Spence, 2011), vowels (Macdermott, 1940), visual shapes (Hanson-Vaux, Crisinel & Spence, 2013; Seo et al., 2010), colour (Gilbert, Martin & Kemp, 1996; Kemp & Gilbert, 1997; Schifferstein & Tanudjaja, 2004) and touch (Demattè, Sanabria, Sugarman & Spence, 2006). Likewise taste has been related to brightness / visual shapes (Deroy & Valentin, 2011; Gal, Wheeler & Shiv, 2011; Spence & Gallace, 2011), colour (O’Mahony, 1983), nonsense words (Gallace, Boschin, & Spence, 2010), auditory tones (Crisinel & Spence, 2009; Rudmin & Cappelli, 1983) and vowel sounds (Simner, Cuskley & Kirby, 2010). Mirroring the relationship between pitch and visual size, auditory pitch has also been related to haptic size (Walker & Smith, 1985). Interestingly, amodal qualities such as shape as determined by either vision or touch, appear to be treated as equivalent during cross-sensory matching (Shields, 2010; Walker, Walker & Francis, 2012). However tactile variations of the 'bouba-kiki' task appear to partially rely on visual experience, as participants with visual impairments have this effect muted (Fryer, Freeman & Pring, 2014). As such even seemingly non-visual correspondences may use visual processes to reinforce their effect. From the perspective of Spence and Deroy (2012), the attenuation of audio-tactile correspondences by visual deprivation may reflect the reduction of correlations between sounds and shapes experienced by a visually deprived individual relative to one with multiple sensory routes to experience this correlation. Similarly, we might expect different environments to reduce this correspondence relative to its extent of correlation in the environment or culture. For correspondences likely based on natural patterns this has typically involved double-checking that many correspondences are universal (Spence, 2011), however some correspondences have been found to exhibit limited variations along cultural lines (Walker, 1987), however the extent to which this is a direct effect of culture is unknown. This leaves open the possibility for correspondences present in infants such as pitch-height, potentially muted or even reversed if the environment only featured high pitched sound sources low in the environment (Spence & Deroy, 2012), similar to unlearning common perceptual assumptions such as the 'light from above' prior in vision (Adams, Graf & Ernst, 2004).

1.2.3. Explaining correspondences

With the ability to match between incredibly wide varieties of sensory properties, any common mechanism underlying these must abstract these properties to a common dimension for comparison. The most commonly cited explanation has been that of matching stimuli based on intensity. At its simplest level, the presence of stimulation in the first modality is more similar to stimulation of a second modality than it is to an absence of stimulation. This works well for explaining loudness-luminance associations as both the auditory and visual cortex increase their neural activation to these stimuli (Goodyear & Menon, 1998; Jäncke, Shah, Posse, Grosse-Ryuken & Müller-Gärtner, 1998). However this is less clear for pitch-luminance associations since pitch is processed tonotopically in the auditory cortex, not through increased activation relative to low pitch (Talavage et al., 2004). One influential account is that of Walsh's (2003) 'A Theory of Magnitude' (AToM), this proposed that sensory features can be abstracted to 'higher' or 'lower' along a sensory dimension in the parietal cortex (Bueti & Walsh, 2009; Cohen Kadosh, Lammertyn & Izard, 2008; Cohen Kadosh, Cohen Kadosh & Henik, 2008; Pinel, Piazza, Le Bihan & Dehaene, 2004). This would mean that sensory features deemed to be 'higher' on each sensory dimension would be treated as more equivalent due to their matching magnitudes relative to their 'lower' counterparts. One difference to 'intensity-matching' is that being higher on a magnitude scale may not necessarily involve the most neural intensity. This can put competing explanations in conflict, for example, a relatively low pitched tone might be 'low' on an abstracted dimension of magnitude however if you were to increase its volume, it would also become high in terms of neural intensity. The neural intensity explanation would expect a loudness-brightness correspondence to occur while the magnitude explanation might expect a colour of low luminance to be chosen. These competing explanations and predictions may help illustrate which best account for these correspondences and furthermore if certain explanations are more common or dominant over others. Individual differences in how correspondences are reached may help to explain why some individuals express loudness-brightness correspondences while others express loudness-darkness correspondences (Marks, 1974). Abstracted theories such as AToM can also be more flexible in what dimension is focused by the individual for the cross-sensory matching. For instance, a complex tactile object could be processed primarily using shape, texture or hardness information, dimensions which have their own unique cross-sensory matches, rather than only how stimulating an object is (Ludwig & Simner, 2013; Servos, Lederman, Wilson & Gati, 2001). In support of applying Walsh's AToM theory to correspondences is that disruption to parietal regions also disrupts interference from incongruent correspondences (Bien, ten Oever, Goebel & Sack, 2011). The finding that abstracted numerosity is topographically arranged in the parietal cortex leaves the door open for specific neurological predictions regarding which correspondences are treated as more or less similar (Harvey, Klein, Petridou & Dumoulin, 2013). For instance, if a specific pitch is associated with an abstract value (and location) in the parietal cortex, its perceived similarity to a specific luminance (which has also been abstracted to the parietal cortex) might be based on the neural proximity of these two abstractions in the parietal cortex. That said, the basis of why certain features (such as pitch) would be treated as 'greater' than others is less well known.

Spence's (2011) theoretical framework for understanding correspondences separates them down to three distinct categories with predictable effects on information processing. The first of these is the 'structural' account where neural connectivity between the corresponding regions leads to enhanced congruent perceptual processing. These regions may either be directly connected, proximal or connected via intermediary regions such as the parietal lobe in magnitude evaluations (Walsh, 2003). One example of this would be loudness-brightness correspondences (Marks, 1974). As this occurs at an early stage of processing prior to conscious evaluations, this would be expected to affect early perceptual processing and influence cognitive decisions (Marks, 1987). The second account is through learned associations through regularities in the environment which are referred to as 'statistical' correspondences. For example, the tendency for smaller animals to have higher pitched vocals than larger animals, leads to a pitch-size correspondence which aides in multisensory integration (Parise & Spence, 2009). Statistical associations can also be artificially brute-forced through trained associations as in attempts to train synaesthesia (Colizoli et al., 2012; Elias et al., 2004; Meier & Rothen, 2009; Rothen et al., 2011). Similar to structural correspondences, these statistical correspondences affect both low-level perceptual processing and higher-level decisions (Evans & Treisman, 2010; Gallace & Spence, 2006). The third account is through shared higher-level meanings, such as through shared words in language or emotional affect. Sharing the terms of "low" and "high" to describe both auditory pitch and spatial location would also lead to a correspondence between the two (Martino & Marks, 1999), while for emotional valence, music in a major chord with a fast tempo was associated with bright yellows, both of which share positive emotional affect linking the two (Palmer et al., 2013). These higher-level links occur at a decisional rather than perceptual level and so would not be predicted to affect behaviours at a pre-cognitive level such as with speeded classification tasks. Contrary to these strict distinctions however is that some correspondences appear to have evidence for occurring at all of these levels. Correspondences between pitch and height occur in pre-linguistic infants under four months of age (Dolscheid, Hunnius, Casasanto & Majid, 2012; Walker et al., 2010). This suggests that minimal (if any) environmental experience is required, which is indicative of a structural explanation. However, Parise, Knorre and Ernst (2014) found that natural auditory scenes in both urban and rural areas featured more high frequency content located higher in space, an effect that is further emphasised by the frequency-filtering properties of the human ear. So pitch-height correspondences are further reinforced through our environment and anatomy, lending support to statistical explanations. These predispositions are likely to influence language with shared terms representing each, resulting in further reinforcement through higher-level semantic correspondences (Martino & Marks, 1999). The potential for one correspondence to occur at multiple levels has many additional implications for their effect on one another. Firstly, it is unclear for correspondences that occur at one or more levels whether this has an impact on the expression of a correspondence, for instance, is a structural correspondence stronger if it is also supported through statistical and higher-level influences? Contrary to this, what occurs if a structural and statistical correspondence are in conflict, is there a hierarchy of influence on perceptual processing tasks? Is it possible to elicit one type of correspondence independent of another? While a variety of explanations can be given to explain a correspondence's origin, the relationship between correspondences has not been explored theoretically.

1.2.4. Correspondences and the present research

Correspondence research offers insights into how even seemingly unrelated phenomena can be processed, represented or abstracted in ways that are treated as similar within the brain. Despite the wealth of variety in correspondence research, some potential correspondences remain unexplored or without a satisfying breakdown of what fundamental sensory characteristics are being linked. For instance, the vowel /e/ is associated with green, even in languages that do not use /e/ in their linguistic term for green (Marks, 1975; Miyahara et al., 2012; Moos, Smith, Miller & Simmons, 2014; Wrembel, 2009). If the explanation is not linguistic, then what attributes of this sound lead to this correspondence? Likewise many explorations of colour correspondences do not control for the influence of other well known correspondences, such as luminance. This makes the breakdown of certain correspondences difficult to specify - for instance do participants choose high pitch for yellow because of its hue, or because it is prototypically brighter (Spence, 2011)? If luminance is controlled for, will the sound-hue mappings remain, and will other new correspondences emerge in these conditions? These questions are explored further in Chapter five.

Correspondences, such as those used in sound-symbolism, have been suggested to be effective in intuitively representing and communicating information about the sensory characteristics of other senses. These sound-visual mappings result in faster and more efficient processing of congruent stimulations. One practical use of this is to communicate visual information to blind individuals using sound in as intuitive a manner as possible. For instance, representing visual height in pitch, horizontal position in panning and visual luminance in loudness, utilises a variety of correspondences to create a new auditory language containing purely visual information. Devices that do this transformation are known as sensory substitution devices (SSDs) and are explored further below. While greyscale vision is routinely translated using existing devices, colour information is also of potential benefit, however to date there has never been a test of whether sound-colour correspondences assist in processing colour information in SSDs. This possibility is explored further in Chapter six.

1.3. Sensory substitution

1.3.1. Introduction

In their pioneering paper, Bach-y-Rita and colleagues reported the production and use of the first 'sensory substitution device' (SSD), a new way of communicating visual information to the blind through their sense of touch (Bach-y-Rita, Collins, Saunders, White & Scadden, 1969). The first SSD, named the tactile-vision substitution system (TVSS) used a bulky television camera to record black and white pictures (see fig. 1.3). These pictures were transformed into spatial patterns of vibration delivered by a 20 by 20 grid of solenoid tappers implanted into the back of a dentist's chair. The spatial position in the image was mapped to the spatial position on the back, with increased luminance felt through increased vibrotactile frequency. So when a user sits in the chair, they can read the 'visual' image from their back. With the user manipulating the camera, a natural feedback loop occurs, where users are able to focus in on different visual areas with a predictable change in tactile stimulation. From this, even congenitally blind users can learn the language of vision, the rules that govern how objects visually change based on visual perspective, distance and shadows. These rules are known as sensorimotor contingencies, for instance, a coin may be circular from the front, a line from the side, smaller in the distance and darker if not exposed to light (Ward & Wright, 2014). Through engaging with the environment using an SSD the experienced location of visual objects begin to feel localised in front of the user rather than as just tactile stimulation on the back (Guarniero, 1974). This sensation can be quite vivid, with Bach-y-Rita (2002) reporting how unexpectedly zooming the camera in on an object, can cause participants to recoil backwards, presumably believing an object to be rushing towards them. This initial proof-of-principle led Bach-y-Rita to suggest that SSDs may help illustrate how neural plasticity would allow the late learning of vision as well as provide practical visual rehabilitation and even visual experiences for the blind (Bach-y-Rita, 1972, 2002; Bach-y-Rita & Kercel, 2003; Bach-y-Rita et al., 1969).

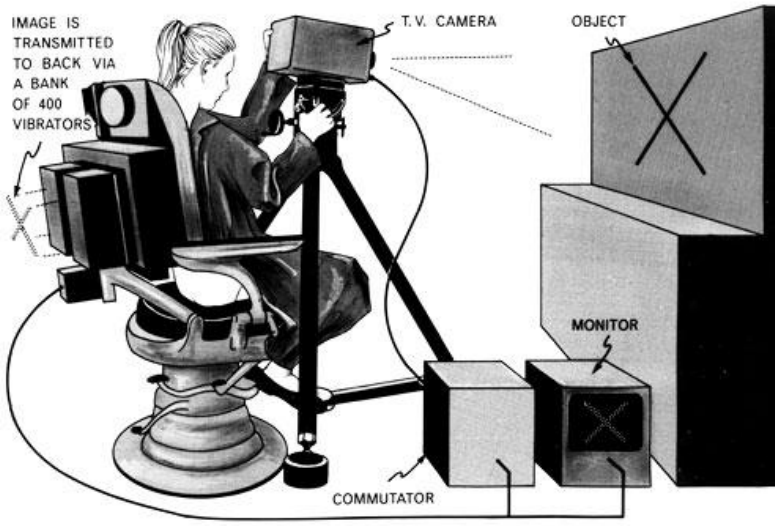


Figure 1.3. Illustration of the tactile-vision substitution system used by Bach-y-Rita. The user controls the T.V. camera to look at objects, with this information being translated into patterns of vibration across a 2D array of vibrators pressed against the user's back. By learning to read off of their back, the user can understand the patterns of luminance in front of them (figure from TCNL, 2014).

Forty-five years have passed since the first paper on SSDs and the field of sensory substitution research has expanded greatly in this time. For visual rehabilitation, multiple approaches to converting visual information into other senses have been tried, assessed through their 'visual' resolution and ability to interact with the environment. The limitations of these, whether through the modality chosen, conversion method or difficulty in interpretation have led to various iterations of these devices. SSDs have allowed us to explore how the brain adapts to process this new type of vision and whether these recruited regions might best be described as a task-dependent rather than inextricably linked to one sense. SSDs also help answer philosophical questions about what it means to 'see,' as well as what is required for a user to report a visual experience.

Subsequent sensory substitution devices were principally concerned with representing three visual dimensions - horizontal position, vertical position and luminance. Encoding these visual dimensions in either tactile or auditory dimensions allows a comparison of how well each modality can interpret the spatial and temporal resolution of vision. These are summarised below.

1.3.2. Vision into touch

The different approaches to encoding 2D greyscale vision in touch have primarily examined which locations of the body which work best for tactile stimulation. Bach-y-Rita's original TVSS used a 20 by 20 grid of tactile stimulators placed 12mm apart, which users could learn on their back. Interestingly there was no loss of spatial localisation when the TVSS was applied to the abdomen, indicating that the conversion of tactile to visual information does not rely on the primary somatosensory representations but the conversion is done at a higher cortical representation (Bach-y-Rita, 1995). The back and abdomen also have a similar two point discrimination threshold, with 9.79mm and 9.78mm respectively, meaning that the spatial distribution of the device (12mm) was comes close to the resolution of the skin (Solomonow, Lyman & Freedy, 1977). This spatial resolution has been sufficient for simple navigation, localisation and interaction with objects; however the resolution is insufficient for cluttered environments (Jansson, 1983). While the spatial resolution of the skin can be further improved through tactile motion (Loomis & Collins, 1978), subsequent tactile-vision SSDs have focused on tactile regions with higher spatial resolutions. One SSD delivered electrotactile stimulation to the fingertip via a 7 by 7 matrix of pins with a 2.54mm spacing between them (Kaczmarek, Tyler & Bach-y-Rita, 1997; see also Kaczmarek, Bach-y-Rita, Tompkins & Webster, 1985). Despite being well over the 1mm discrimination spacing capable for pressure (Van Boven, & Johnson, 1994), discriminating geometric patterns through electrotactile stimulation appeared to lose some resolution relative to raised dot patterns. Users were able to do simple form recognition at the highest comfortable currents; however discrimination for lower intensities was impaired, limiting the range of intensities usable to represent luminance information (Kaczmarek & Haase, 2003). This device formed the precursor to the Brainport, which provided similar electrotactile stimulation to the tongue via a 'tongue display unit' (TDU - see fig. 1.4). The tongue itself is an area that provides both a higher tactile resolution and increased electrical conductivity through the tongue's saliva (Bach-y-Rita, Kaczmarek, Tyler & Garcia-Lara, 1998; Van Boven, & Johnson, 1994). This resulted in further improvements to form perception while only requiring 3% of the current used in the fingertip version. The TDU has been used for obstacle avoidance during navigation in the congenitally blind as well as recruit similar occipital regions to those used in visual navigation (Chebat, Schneider, Kupers & Ptito, 2011; Kaiser, 2014; Kupers, Chebat, Madsen, Paulson & Ptito, 2010). Due to the simultaneous input of multiple spatial points, the TDU is also well suited for motion discrimination, an activity that activates 'visual' motion areas in sighted and the congenitally blind (Matteau, Kupers, Ricciardi, Pietrini & Ptito, 2010).

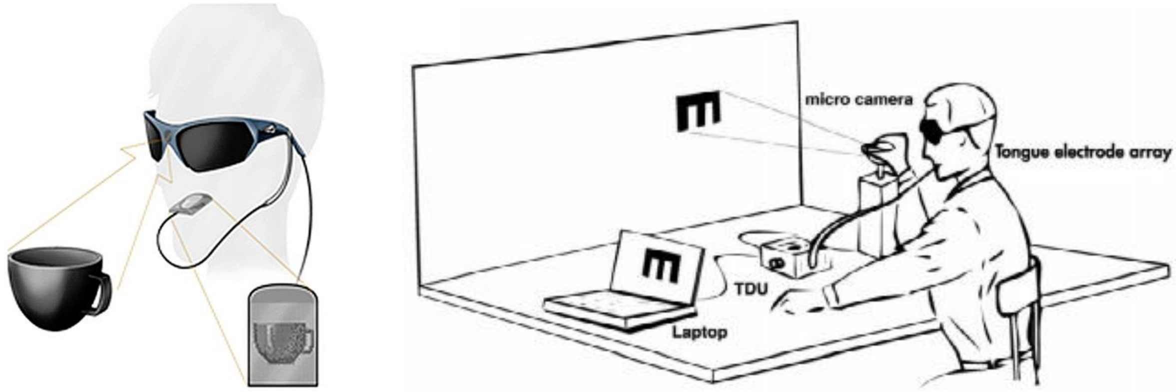


Figure 1.4. Brainport (left) and Tongue Display Unit (right). 2D greyscale images are picked up by a webcam and translated into a pattern of spatial electrical stimulation delivered to the tongue (figures from Lee et al., 2014 and Sampaio, Maris & Bach-y-Rita, 2001).

1.3.3. Vision into sound

While spatial information in tactile-vision SSDs have typically taken advantage of directly mapping to the skin's surface area, converting fine-grained spatial information into sound has produced more varied design solutions. In 1992, Meijer introduced the 'vOICe' (middle letters sounding "Oh I See"), a visual to auditory SSD that converted 2D greyscale video to variations of pitch, loudness, panning and time (see fig. 1.5). Using the default settings, the vOICe takes a 64 by 64 pixel image with 16 levels of luminance every second from a live camera feed. The vertical position of a pixel denotes its pitch, while the luminance of the pixel gives us its loudness. The image is fed to the user one column at a time starting on the left side of space before scanning to the right side of the image. During this the sound also pans from the left to right ear to further assist in horizontal localisation. The combination of all of these factors into a single sound is called a 'soundscape.' As such, a bright object in the top left of space would be heard as loud, high pitched tone in the left ear at the beginning of the soundscape. With training the vOICe has proven to be effective at the localisation, recognition and discrimination of objects (Auvray, Hanneton & O'Regan, 2007; Proulx, Stoerig, Ludowig & Knoll, 2008). Subjective reports suggest that localisation is felt as similar to vision, while object recognition feels more like hearing, most likely a result of the increased focus on small variations in the auditory signature of different objects (Auvray et al., 2007). As such it could be suggested that learning objects could be done through memorising an objects' auditory signature, however as users gain experience with the vOICe's conversion rules, these abilities can be increasingly transferred to discriminating novel objects (Kim & Zatorre, 2008). As such a users' initial higher-level understanding of the conversion rules limits their discrimination ability in a way that engagement with the lower-level processes of loudness, pitch and timing discrimination does not. Expertise with the vOICe by blind users has been associated with visual sensations not only to vOICe signals, but automatically to tones within the vOICe's range that do not contain visual information

(Ward & Meijer, 2010). Expert users of other SSDs which turn colour information into pure tones, have also been reported to automatically mis-interpret sounds that do not contain visual information. For instance, colour-blind SSD expert user NH reports that the sounds of ‘jingling keys’ automatically elicits colour associations (NH, personal communication, August 29th, 2012). These lines of evidence in combination suggest that progression to expertise occurs alongside a progression from a higher-level conceptual understanding of an SSD’s conversion rules to utilising lower-level perceptual processes to automatically interpret direct sensory stimulation. It is currently unclear if this progression from higher-level to lower-level processes would be similar for SSDs that do not utilise correspondences known to operate at a perceptual level (Spence, 2011). Alternatively it could be considered that some SSDs would always maintain higher-level or intermediary processes similar to the operation of certain correspondences (e.g. sound-colour correspondences mediated by emotion – Palmer et al., 2013) or direct / indirect pathways used in certain types of synaesthesia (Bargary & Mitchell, 2008).

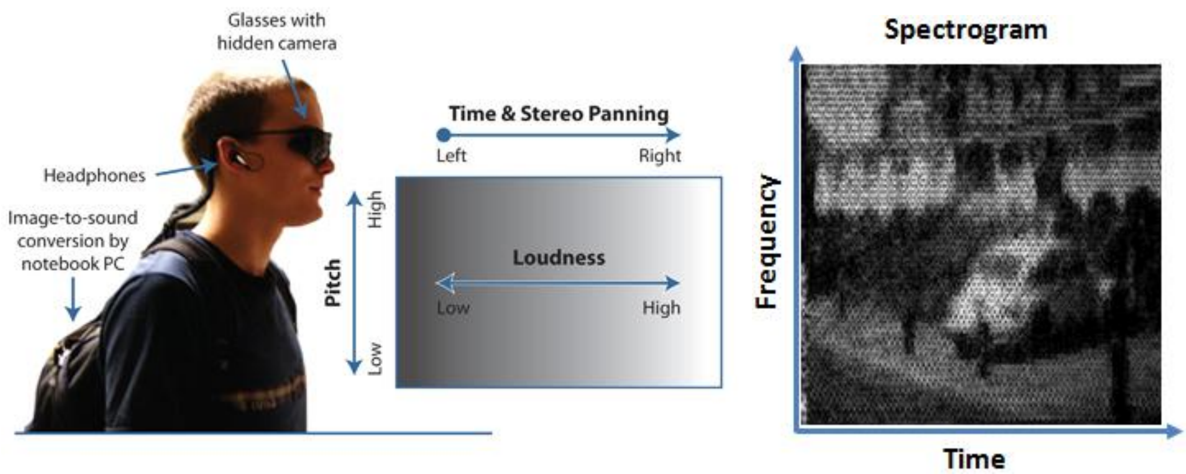


Figure 1.5. 'vOICe' sensory substitution device (left) and spectrogram output (right). A greyscale image is picked up via a head mounted webcam, and is translated into sound. Vertical height is translated into pitch, luminance into loudness and each column is fed piecemeal to the user over time. There is a high retention of information as indicated by spectrogram outputs (figures from Proulx et al., 2008 and Meijer, 2014).

Although the vOICe has been highly influential, there have been multiple approaches to representing greyscale vision through sound. Similar to the vOICe, the 'Prosthesis for Substitution of Vision by Audition' (PSVA) is an SSD that also encodes vertical position in pitch and luminance in loudness, however the horizontal position is also encoded in pitch as well as binaural intensity (Capelle, Trullemans, Arno & Veraart, 1998). The visual image is segmented into an 8 by 8 grid of large pixels with another 8 by 8 grid of smaller pixels at the centre. This creates higher resolutions at the centre, reminiscent of the fovea. Participants have effectively used the PSVA in pattern discrimination, object localisation and recognition tasks, while congenitally blind individuals have learnt visual depth cues (Arno, Capelle, Wanet-Defalque, Catalan-Ahumada & Veraart, 1999; Arno et al., 2001; Poirier, De Volder, Tranduy & Scheiber, 2007; Renier & De Volder, 2010; Renier et al., 2005a). This proficiency also extends to experiencing illusions based on depth and vertical-horizontal comparisons through sound for those with prior visual experience (Renier, Bruyer & De Volder, 2006; Renier et al., 2005). Behaviourally, it seems that processing the SSD signal seems to also retain visual biases.

Beyond the PSVA's 'fovea,' other attempts at visual pre-processing have also been tried. Cronly-Dillon, Persaud and Gregory (1999) introduced the SmartSight SSD, which gives users the ability to solely select certain areas or elements of the visual image for sonification. The method for conversion into sound is similar to the vOICe, with columns presented in a piecemeal fashion to the user over time, with pitch-height and luminance-loudness mappings. The vertical axis is divided into 50 Pianola notes. This has been used to deconstruct and recognise both simple shapes and complex natural images in both blindfolded and late blind individuals (Cronly-Dillon, Persaud & Blore, 2000; Cronly-Dillon, Persaud & Gregory, 1999).

One final auditory conversion device translating 2D greyscale images is the Vibe (Auvray, Hanneton, Lenay & O'Regan, 2005). Several 'sound sources' are placed on the visual image, each of which has a receptive field that sums the luminosity of nearby pixels. This average luminosity is turned into the loudness of the sound source, with its spatial position determined vertically by pitch and horizontally by the balance between stereo channels. All of the sound sources are played simultaneously so temporal resolution remains good and the spatial resolution can be varied by the number of sound sources and their receptive fields (Hanneton, Auvray & Durette, 2010). Preliminary testing using the vibe has examined the effect of interaction with the environment on distal attribution of stimuli as well as navigation within a natural setting (Auvray et al., 2005; Durette, Louveton, Alleysson & Hérault, 2008).

1.3.4. The resolution of SSDs

Through the translation of visual information to the user there are several bottlenecks that constrain the preservation of information. The first is the device itself, from the visual resolution to the translation procedure. Next is the point of contact between the SSD and the user. Since the informational bandwidth of the eyes is far greater than the ears or fingertips (Jacobson, 1950, 1951; Kokjer 1987), visual information has to sacrifice resolution, be it temporal, spatial or qualitative like colour. The last bottleneck is the users' ability to discriminate and interpret signals, giving us the devices' actual acuity.

One way of comparing vision to SSDs is through assessing their visual acuity. The most common test used for this is the Snellen 'E'-orientation test, where the letter 'E' is rotated either 0, 90, 180 or 270 degrees and users discriminate which of these four orientations it is. The final measure of visual acuity is given in how close (in feet) the stimulus needs to be to discriminate (the numerator) compared to the distance someone with normal vision could discriminate it (the denominator). So 20/20 would be average while 20/200 reaches the legal definition of blindness within Europe and the USA. Since SSDs use a camera they can also vary in their field of view (FOV). Since you could 'zoom-in' to increase acuity, the fixed FOV is reported here also. An FOV of under 20° is also considered legal blindness irrespective of acuity within that FOV.

The only tested tactile-visual SSD is the TDU with a 12x12 matrix, reaching an initial resolution of 20/860 and increased to 20/430 with training to near 100% accuracy with a 54° horizontal FOV (Sampaio, Maris & Bach-y-Rita, 2001). Chebat, Rainville, Kupers and Ptito (2007) tested a smaller 10x10 matrix, where the resolution dropped to 20/1800-8400 for 70% correct with a lower 29° FOV, where blind individuals reached the highest resolutions. Both training and blindness appear to increase acuity beyond the expected resolution for discriminating two individual pixels with the device worked out to be 20/2160 for the 12x12 TDU at a 54° FOV (Haigh, Brown, Meijer & Proulx, 2013). So how can users exceed this? One way is through using movement, judging the movement required for a stimulus to 'jump' to the next pixel. Alternatively, variations in greyscale can also allow users to infer spatial information, for instance, if an SSD user knows a stimulus is white and the background is black, then if this stimulus is smaller than the FOV for each ‘pixel’ of spatial resolution, then the stimulus only partially contributes to that pixels' luminance. So a white stimulus at half the FOV of each pixel would be delivered to the user as 50% luminance (or grey) from which they can infer the stimulus is half the size of a pixel indicating 100% luminance.

Acuity for auditory SSDs have only used the vOICe, with passive screen reading of the Snellen test reaching up to 20/1882, while active explorations using a webcam reach up to 20/737 at a 75% correct rate with 66° FOV (Haigh et al., 2013). In line with this Striem-Amit, Guendelman and Amedi (2012) report that highly trained early blind users were able to reach between 20/200-600 with a 66° FOV for the Snellen task at a rate of 60% correct. There are additional advantages for active explorations with the vOICe, since users can represent the 'E' in high or low pitches by raising, lowering or tilting the camera, users can work within the pitches they find easiest. However the different representations for horizontal and vertical axes makes certain orientations easier to discriminate than others (Haigh et al., 2013), for instance, vertical height is judged through a comparison of the lowest and highest frequencies played, while horizontal length is judged through the length of time a frequency is played. Haigh et al (2013) also found that participants with musical training performed better with the vOICe, suggesting that the sensory bottleneck of discriminating pitch accurately was a barrier for those without musical experience (Micheyl, Delhommeau, Perrot & Oxenham, 2006). This has been suggested previously with individual musicians outperforming non-musicians in object recognition tasks with the vOICe (Auvray et al., 2007). From this, it would also be expected for early-blind individuals to perform better due to enhanced discrimination of pitch irrespective of musical experience (Wan, Wood, Reutens & Wilson, 2010). In combination these help explain the high resolution found for congenitally blind vOICe users relative to blindfolded controls (Haigh et al., 2013; Striem-Amit et al., 2012).

1.3.5. Visual phenomenology

SSDs provide blind users with the ability to use and discriminate the sensorimotor 'rules of vision' using touch and hearing (Ward & Wright, 2014). This extends an ability to test amodal theories of consciousness such as sensorimotor theory (O'Regan & Noë, 2001). This theory puts forward that visual phenomenology is dictated by our interaction with the environment and mastering the sensorimotor rules that define this interaction. From this, it follows that the eyes are not an inherently privileged source of information, but that experience and mastery of the rules of vision from any modality should produce visual consciousness. Ward and Meijer (2010) published interviews with two highly experienced late blind users of the vOICe SSD. The accounts of late blind individuals are particularly important because they have the ability to compare their current sensory substitution experiences with their memories of vision. Both users PF and CC had decades of visual experience prior to their causes of blindness, and after over 20 years for PF and 10 years for CC of being registered blind started to use the vOICe SSD. Both users have minimal residual light perception, and after 10 years experience with the SSD, PF has described the vOICe's soundscape as generating luminance in the same location as her residual vision. CC also notes that veridical and SSD perceptions of the same object exist in the same spatial location. This mirrors behavioural work on novice SSD users having their spatial sense influenced by cues from SSDs to influence both eye movements and visually directed behaviours (Levy-Tzedek et al., 2012a, 2012b; Wright & Ward, 2014). Users PF and CC also report slowly learning depth information, smooth motion and colour beyond what is directly coded in the vOICe (Ward & Meijer, 2010). These factors can be both extrapolated from consistent interactions between the user, environment and SSD signal to make predictions in accordance with sensorimotor rules, or with the rules from visual memory (e.g. Christmas trees are green) rather than the environment (O'Regan, 2011).

Visual experiences from tactile interfaces are less well substantiated, with one early blind user of the TVSS using visual terminology related to perspective, observation and spatial location but noting that they use visual terms for lack of more accurate ones (Guarniero, 1974). Seemingly contrary to visually-specific sensorimotor rules being necessary for the conscious experience of vision are reports of visual luminance in response to non-SSD tactile stimulation. Ortiz et al (2011) report that tactile orientation discrimination tasks could over the course of three months result in visual experiences of luminance for blind users with minimal or previous experiences of sight. This coincided with increased visual cortex involvement in blind individuals. Similar tactile tasks have been known to recruit visual orientation regions so this may play a role in the reported visual imagery (Sathian & Zangaladze, 2002; Zhang et al., 2005).

1.3.6. Visual processing

The recruitment of visual areas for SSDs showcases the brains' plasticity for processing a variety of stimuli. Both auditory and tactile SSDs have been found to elicit early visual cortex activations in trained blind users or secondary visual associative areas in blindfolded users (Arno et al., 2001; De Volder et al., 1999; Pollok, Schnitzler, Stoerig, Mierdorf & Schnitzler, 2005; Ptito, Moesgaard, Gjedde & Kupers, 2005; Renier et al., 2005). This appears to serve a functional role with disruption of the occipital cortex leading to impairments in using the vOICe in blind expert users (Merabet et al., 2009). Other regions for processing shapes, words and bodies have also been found to be activated by the discrimination of each using the vOICe (Amedi et al., 2007; Streim-Amit & Amedi, 2014; Striem-Amit, Cohen, Dehaene & Amedi, 2012). The ability to recruit these brain regions suggests that they not inherently linked to one modality and operate according to the general discrimination demands (Proulx, Brown, Pasqualotto & Meijer, 2014; Reich, Maidenbaum & Amedi, 2012). Ricciardi, Bonino, Pellegrini and Pietrini (2014) suggest that visual experience is not a prerequisite to developing systems for form, space and movement but instead vision plays a role in refining these relationships. As vision is important for representing external space (Pasqualotto & Proulx, 2012), SSDs provide an avenue to teach spatial relationships between objects and observers early in development as well as aid multisensory learning (Miletic, Hughes & Bach-y-Rita, 1988; Proulx et al., 2014).

1.3.7. Greyscale and colour vision

The most widely reported studies using sensory substitution have focused on the conversion of greyscale images into touch and sound. As a result, studies on visual resolution, illusions, object discrimination, recognition and navigation using the SSDs mentioned here have used highly simplistic and artificial high contrast stimuli (Auvray et al., 2007; Bach-y-Rita et al., 1969; Chebat et al., 2007; Haigh et al., 2013; Sampaio, Maris & Bach-y-Rita, 2001; Striem-Amit et al., 2012). However luminance information alone at the low spatial resolutions seen by users may not be the optimal solution for more natural environments or stimuli. Since natural environments feature large changes in environmental illumination, shadows and contrast, edges defined by luminance alone are likely to be unreliable. Other properties of colour such as hue remain more resistant to changes in illumination, this along with saturation information provides multiple dimensions can not only define object edges but also provide clues to an object's identity. Chapter two looks at the utility of colour information for sensory substitution in more depth, alongside examining the wide variety of ways that both tactile and auditory SSDs have tackled the issue of colour representation. Experiments using these colour SSDs on navigation, object recognition, colour knowledge and resistance to changes in environmental illumination are reported, illustrating some of the advantages colour representation can have for the end user. Chapter two finishes with a look into promising future directions for combining new technology as well as our knowledge of tactile / auditory discrimination thresholds and correspondences for optimal representations of colour in SSDs.

1.4. Colour definitions, transformations and rationales

1.4.1. Introduction and definitions

In describing a particular colour, many terms can be used to both describe its physical characteristics as well as its perceptual characteristics; the most common terms are luminance, brightness and lightness that chart a progression from dark to light, while hue, saturation and chroma refer to deviations away from greyscale values. Luminance typically refers to the intensity of light emitted from a surface in candelas per square metre or cd/m^2. It can be used in colour models where the actual luminance cannot be known (e.g. HSL's 'L' luminance dimension) or in colour models that account for the intensity of outputted light (CIE XYZ's 'Y' luminance dimension) and human observer's perceptual characteristics (e.g. CIE LUV's 'L' lightness dimension). Brightness is the subjective measure of luminance intensity whereby one source appears to emit either more or less light relative to another. While luminance increases are proportional to physical power, brightness however, is a function of the sensitivity of human vision's response characteristics and is as a result non-linear in relation to luminance increases. Lightness is an attempt to linearize brightness estimates through a modified cube root of luminance values. This method can be found in L values for CIE LUV, LCh_{uv} and LAB colour spaces. Hue is when a colour appears to be similar to red, yellow, green, blue or a combination of two of these in response to different spectral power distributions of a given stimulus. Saturation is the observed 'colourfulness' of a given area relative to its brightness in response to dominant wavelengths of a stimulus' spectral power distribution. Related to this is chroma, the subjective intensity or purity of colour in a stimulus (Poynton, 1997).

1.4.2. Colour spaces

A single specific shade of colour can be represented in a multitude of ways as a numeric representation within a colour space. Each of these different colour spaces have different assumptions and features intrinsic to their space and utilise different dimensions to navigate to a given colour. Limitations on the amount of colours that can be represented in a given colour space are usually a constraint of any output device rather than the space itself. Below is an account of a variety of colour spaces mentioned in the thesis, including information on their purpose, rationale, relation to one another and mathematical transformations used in analyses undertaken in this thesis.

1.4.3. RGB space

RGB space is an additive colour space based on combining the intensities of three different chromacities of light relating to red, green and blue. The precise spectral qualities of these R, G and B channels are intrinsic to the hardware displaying them and their settings. Each of these light sources are controlled and represented by the R, G and B dimensions that make up RGB space, which each dimension ranging between 0 and 255. The combination of all three can be superimposed through either the trichromatic output reaching a diffuser, using high enough temporal frequencies to remove flicker or through combinations of light sources too small for the observer to spatially resolve. Increasing each value separately increases both the saturation and luminance of that individual colour, varieties of mixing these together produce different hues (e.g. high red and green values produce a final colour with a yellow hue), whereas a combination of equal R, G and B values produce de-saturated colours ranging from black through grey to white depending on the level of luminance from all channels combined (see fig. 1.6). Each of the dimensions has a potential for 256 different levels of output due to the limitations of information that an 8-bit byte can represent. In combination, this allows for potentially 16777216 different shades of colour.

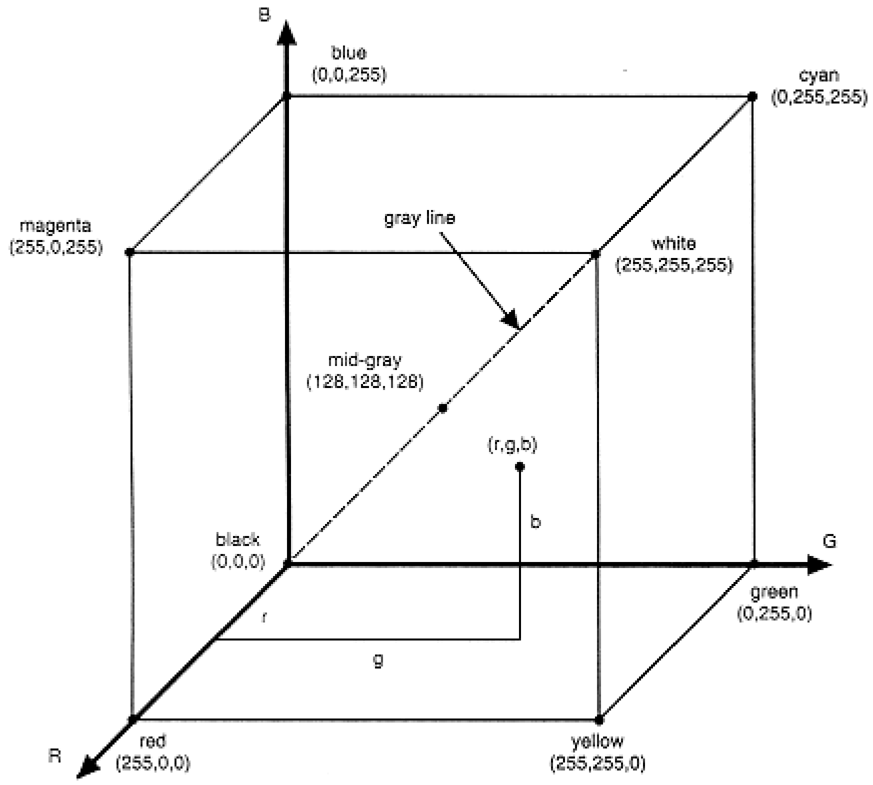


Figure 1.6. RGB Space. Values for each of the R, G and B dimensions can vary between 0-255, in combination these produce unique colours varying in hue, saturation and luminance (Parke, 2003).

1.4.4. Criticisms of RGB space

While the digital representations of a given colour in RGB space may be the same across devices, this space does not account for variations in monitor output for each of the R, G and B channels, therefore it is considered a device-dependent colour space. As such, the same RGB co-ordinates may produce different colours depending on the monitor the colour is displayed on. An attempt to counter this inter-device variability was produced using standardised RGB space (sRGB) whereby the intensities of the R, G and B channels are assumed to be similar to a standardised factory output. This assumption of sRGB space is presumed in transformations from RGB space to other colour spaces in order to minimise error in the final colours produced after transformation, alternatively certain colour spaces (e.g. CIE XYZ, LUV) can account for objectively measured RGB values. RGB space also does not attempt to create a perceptually uniform space, so distances travelled within RGB space between two colours does not relate to perceptual distances between these colours to a human observer. Each of these shortcomings has been reduced through the use of colour spaces generated by the International Commission on Illumination (CIE). It is possible to convert RGB values to CIE colour spaces through knowing the intensities of red, green and blue chromatic light. Since CIE colour spaces can account for the specific phosphor intensities of the R, G and B channels on output devices, CIE colour spaces can create device-independent colour spaces to produce objectively similar colours across multiple devices. CIE colour spaces also take into account the cone cell response characteristics of human observers in order to produce a large variety of available colours (e.g. CIE XYZ). This distribution of colours can be put into a perceptually uniform arrangement (e.g. CIE LUV), which means that distances travelled within these spaces relate to perceptual colour distances as judged by human observers (Hunt & Pointer, 2011).

1.4.5. Cone responsivity characteristics and CIE 1931 colour space

In perceiving colour, the eye has three types of cone cell that each physiologically respond to a certain range of wavelengths while maximally responding to one particular wavelength. These cone cells respond preferentially to short (*S*), middle (*M*) or long (*L*) wavelengths. The peaks in spectral sensitivity are between 420-440nm for *S* cone cells, 530-540nm for *M* cone cells and 560-580nm for *L* cone cells. While the peaks of these sensitivities do not overlap, the total range of wavelengths that each cone is responsive to is far wider, overlapping between cones. This means that peak wavelengths for *M* cones would also stimulate *L* cones (see fig. 1.7). Hypothetically, to reach a certain response from a specific cone cell but no response from the others (e.g. Maximal *M* cone response and no *S* or *L* cone response) would require negative values from the other cones.

However this is impossible to produce in additive colour models. To create a space where the maximal variations of cone responses can be elicited while retaining only positive values to generate a colour, a new set of functions need to be created. These functions need to account for metamerism, where two colours can appear the same to an oberserver and yet consist of different spectral power distributions. These insensitivities of the observer correspond roughly to the LMS curves (see fig. 1.7) and can be used to base outputs of a device to correspond to specific colours. Colour spaces can seek to elicit a maximal variety of perceptual colours on an output device through utilising this information and proportions of red, green and blue stimulation to the observer as seen in CIE's 1931 XYZ colour space (Hunt & Pointer, 2011).

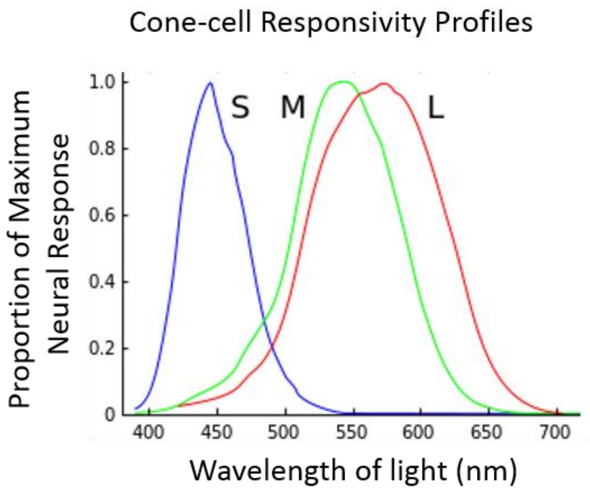


Figure 1.7. Cone cell spectral responsivity profiles for short, middle and long wavelength responsive cones (Adapted from Fairchild, 2005).

CIE XYZ colour space is a tristimulus additive colour model which sort to implement data from colour-matching experiments with human observers to create a colour space with a wide gamut of available colours and no negative values that would serve as a basis for a variety of future standardised colour spaces. The X, Y and Z stimulus values represent the contribution of R, G and B phosphors to the responsivity of *L*, *M* and *S* cones. For example, equal powers of red, green or blue light are not perceived as equally bright, with green light perceived as the brightest. As such, the *M* cone response profile provides the greatest contribution to overall brightness and provides the greatest influence on the CIE Y dimension which approximates to the luminance of a colour. The CIE Z and X dimensions contain all possible chromatic information at a specific luminance (as denoted by the CIE Y value). CIE Z has a closer approximation to the *S* cone response profile receiving a majority of influence from blue phosphors. The CIE X dimension uses a profile that combines multiple cone responsivity characteristics and fits these so that no negative numbers can be produced. This means that CIE XYZ can be produced from R, G and B phosphers in an additive fashion. The values used to determine this space relate to data obtained through human observers in colour matching experiments for a specific visual angle (normally a 2 degree visual angle) called the CIE 1931 2° Standard Observer.

1.4.6. CIE 1931 2° standard observer and XYZ tristimulus values

As cone density varies throughout the retina, the visual angle that a colour stimulus spans is important in testing equivalent regions within the same participant, especially in colour matching experiments. A visual angle of 2° is typical and used in the CIE 1931 Standard Colourimetric Observer. This involves perceptually matching a white stimulus comprised of all visual wavelengths to a colour generated from contributions of red, green and blue light sources (i.e. 3 wavelengths). Matching these two perceptually similar colours that are comprised of different spectral distributions is known as metamerism. The red, green and blue light sources are not added in an equal mixture to create an equivalent white. This is because there is an unequal distribution of red (*L*), green (*M*) and blue (*S*) responsive cones. The contribution of chromatic lights to a final brightness measurement is unequal, with green providing the highest contribution to overall brightness and blue providing the lowest. Thus in combining red, green and blue light sources to create an equivalent brightness to a white light, blue light sources have a lower luminance value and perceptual contribution than red, while green light sources have the highest values and perceptual contributions as measured in objective luminance values of candela per square metre (cd/m^2) (Wright, 1928; Guild, 1931). Red light (700 nm) results in 1 cd/m^2, green light (546.1 nm) is 4.59 cd/m^2 and blue light (435.8 nm) is 0.06 cd/m^2. These absolute values become an assumed luminance output for the R, G and B channels in XYZ colour space, with the previously stated luminance cd/m^2 values now given a power of 1 to denote parity with these original luminance ratings when RGB space is transformed into CIE XYZ. However, unless a monitor has a standardised RGB output (and factory settings), it is unlikely that a given monitor would output values of precisely this luminance. As a result this power value needs to be calibrated to reflect the monitor’s actual output and mimic the original mixtures of R, G and B used for brightness matching. Any deviation of luminance output for the red, green or blue channels from these original measurements can be accounted for by changing the power values during colour space transformation so that specific luminances can be reliably calculated for output on a monitor (Hunt & Pointer, 2011).

1.4.7. Gamma correction for luminance linearity

In calibrating for the monitor’s luminance output for the R, G and B channels, an assumption of luminance linearity is present for each channel. This means that between 0 and 255 for each colour channel there should be an equal increase in luminance for each increase in colour value. This assumption can be checked using a colorimeter and deviations from a linear increase can be corrected through replacing old R, G or B values with new values that produce the appropriate level of luminance output. This process involves measuring the luminance output from the R, G and B channels separately across the 0 to 255 range and describing this relationship using a function. Deviations between this function and a linear relationship between colour values of 0 and 255 are established and a reverse of this function is fitted to the R, G and B channels to establish replacement colour values. By using the replacement colour values, the colour output on the monitor can now increase in a linear fashion from 0 to 255. This process is illustrated with data from Chapter 5 in fig 1.8.

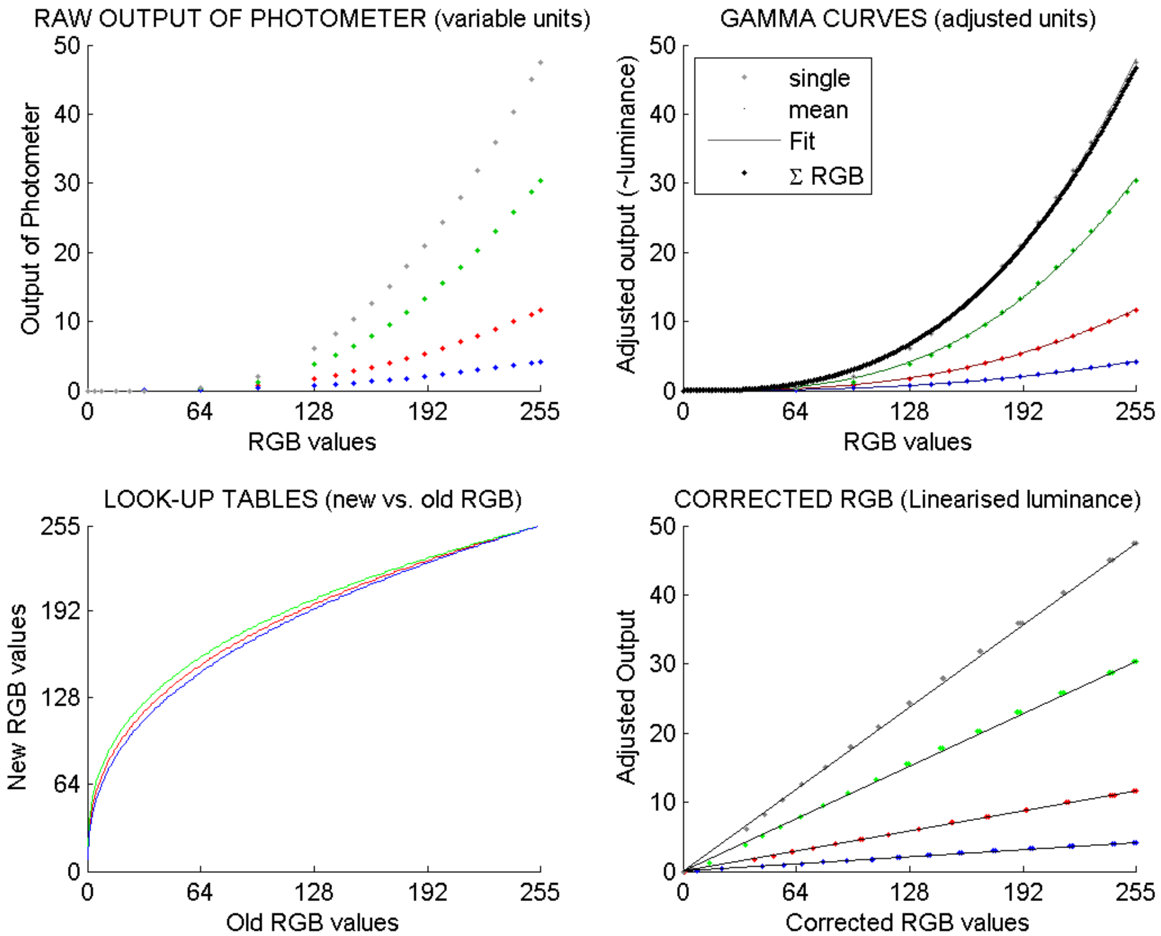


Figure 1.8. Luminance measurements and linearity corrections for R, G and B channels for the CRT monitor used in Chapter 6. Top left graph shows luminance measurements from a ColorCAL MKII Colorimeter from Cambridge Research Systems for individual R, G, B and all channels for the range of 0 to 255 as indicated by the congruently coloured dots for each channel. Top right graph shows a function being fitted to estimate the remaining points not measured by the colorimeter. This results in a gamma curve. Bottom left graph shows the inverse deviation from a straight line between minimal and maximum R, G and B values that the gamma curve has for each channel. This is done in order to gather the replacement R, G and B values – for instance, an old B value of 64 is replaced with a new value of 146. This replacement information is contained within a 'lookup table' (LUT). Bottom right, the result of replacing all old values with new values from the LUT is a linear increase in outputted luminance for the R, G and B channels during an increase from 0 to 255.

1.4.8. Negative values in additive colour models

Certain wavelengths of colour can stimulate multiple cones to different degrees (see fig. 1.7), so for example, a wavelength of 546.1nm will preferentially stimulate not only green-sensitive cones but red-sensitive cones as well. As such, it is not always possible to individually stimulate a specific cone, so strictly additive models of colour such as RGB space will need to take into account these relative contributions of a given wavelength to all cones using spectral weighting functions. Using RGB directly runs into the problem of having to use hypothetical negative values, for instance, if green light (546.1nm) produces a response profile from *L* cones (red) of 75 units of power, *M* cones (green) of 100 units of power, and *S* cones (blue) of 1 unit of power. Then to elicit a green with a reduced response from red-sensitive cones, hypothetically a negative red value is required or this value would need to be added to the stimulus being matched to, rather than the mixture (see fig. 1.9). However, since only positive values can be used in adding light mixtures together, then the three dimensions used to represent colour instead each have contributions of R, G and B channels in them as used in CIE XYZ space.

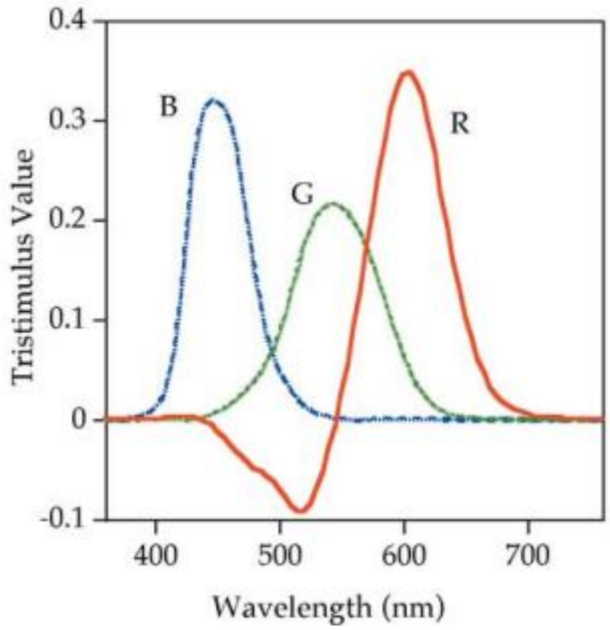


Figure 1.9. Spectral tristimulus values for monochromatic primaries at 435.8, 546.1 and 700nm (Fairchild, 2005).

1.4.9. RGB to CIE XYZ transformations

In 1931, CIE created a colour space that could be calculated from objective measures of a device's luminance output to produce a system that produces a wide variety of colours, contains no negative values as well as being device-independent and calculated from spectral power data from brightness matching experiments (Wright, 1928; Guild, 1931). The standardised dimensions of R, G and B can be translated into perceptually meaningful tristimulus values, CIE X, Y and Z, worked out using the following formulae:

$R_{proportion} = R / 255$

$G_{proportion} = G / 255$

$B_{proportion} = B / 255$

If the proportion of R, G or B is over 0.04045 the following calculation is done for that specific channel:

$R = ((R_{proportion} + 0.055) / 1.055)^{2.4}$

$G = ((G_{proportion} + 0.055) / 1.055)^{2.4}$

$B = ((B_{proportion} + 0.055) / 1.055)^{2.4}$

If alternatively the proportion of R, G or B is under 0.04045 then the following calculation is done for that specific channel:

$R = R_{proportion} / 12.92$

$G = G_{proportion} / 12.92$

$B = B_{proportion} / 12.92$

These values are then multiplied by 100 and then provide the values which are weighted in XYZ space according to data from the CIE standard colorimetric observer (in this case 2° visual angle during daylight conditions, referred to as D65).

$R = R * 100$

$G = G * 100$

$B = B * 100$

$X = (0.4124 * R) + (0.3576 * G) + (0.1805 * B)$

$Y = (0.2126 * R) + (0.7152 * G) + (0.0722 * B)$

$Z = (0.0193 * R) + (0.1192 * G) + (0.9505 * B)$

This calculation is done for XYZ space in Chapter 3, 4 and 6 during the process of transforming to CIE LUV space. While conversions to a perceptually uniform space such as CIE LUV require the XYZ values to reflect the monitor's calibration to be entirely accurate, even considering the error introduced through a non-calibrated monitor, these spaces provide advantages to both the sensitivity and specificity in establishing colour consistency in synaesthesia under non-calibrated conditions over other colour spaces (Rothen et al., 2013). For the colour corrected CRT monitor used in Chapter 5 an alternative weighting is provided below as a result of luminance and chromaticity measurements taken for maximum R, G and B channel values. Relative proportions of X, Y or Z to X+Y+Z give chromaticity co-ordinates designated by lower-case x and y dimensions. These x and y dimensions are designed to be normalised in that they no longer contain brightness information, so travelling along a particular hue will also vary its brightness (see fig. 1.10). The output of the ColorCal colorimeter gives readings for CIE xyY space, where x and y give chromaticity values from 0 to 1, while Y gives a luminance value out of 100 in cd/m^2.

Table 1.2. Measured CIE xyY locations for Red, Green and Blue channels (from Chapter 5).

CIE xyY location	Red	Green	Blue
x	0.627	0.283	0.151
y	0.343	0.615	0.069
Y	11.601	30.346	4.210

These locations can be plotted on chromaticity diagram to provide an approximate gamut of available colours from the CRT monitor using the location of maximum R, G or B phosphor values for the monitor in CIE xyY space. In plotting this 3D space on a 2D graph, luminance information (Y) is omitted in favour of the CIE x and y co-ordinates to retain saturation and hue information. This is typically displayed as three points of a triangle with each point relating to maximal R, G and B values. The space between these points gives an indication of the available gamut of colours available to a given output device (see fig. 1.10).

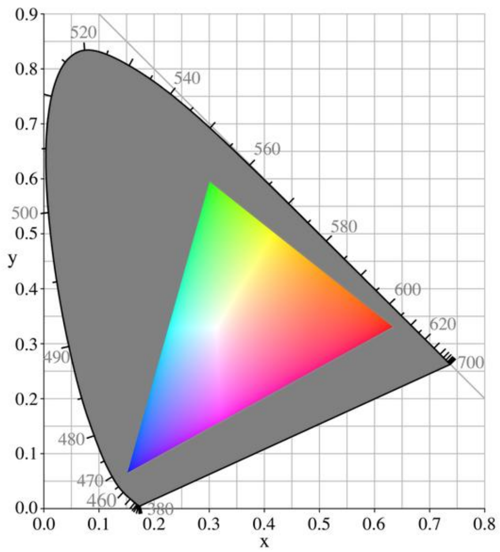


Figure 1.10. Example of a gamut in CIE xy space. The chromaticies for maximum R, G and B values plotted as chromatic co-ordinates in CIE xy space, the space between these three points correspondences to the total available range of colours that can be output on a given device. The grey bounds represent possible colours that cannot be achieved as they are outside of the monitor's gamut. The outer bounds refer to the maximal sensitivity expected to short, middle and long wavelength sensitive cones in the human observer responding to different single wavelengths between 380 and 700nm. Luminance information is not presented.

Now that the chromaticities of the primary channels are known, the R, G and B channels can be set to yield an objective output consistent with previously measured CIE xyY values for a defined colourimetric observer. In Chapter 5, a standard illuminant C with xyY values of 0.3101, 0.3162 and 50 (Fairchild, 2005) was chosen as the best approximation of measured monitor output for RGB values of 255, 219 and 209, producing mean CIE xyY values of 0.309 (SD = 0), 0.316 (SD = 0) and 51.611 (SD = 0.042) across six measurements. Transformation of CIE xyY into CIE XYZ values can be done using the following formula:

CIE X = Y * x / y

CIE Z = Y * (1 – x – y) / y

For a maximum luminance of Y = 100, this yields XYZ values of 98.07, 100 and 118.18 which act as reference values in transforming into CIE LUV and LCh_{uv} colour space.

1.4.10. CIE XYZ to CIE LUV & LCh_{uv} transformations

CIE LUV space is a transformation of the perceptually meaningful CIE XYZ values into a perceptually uniform colour space where spatial distances between colours relate to their percieved colour differences. The dimensions are primarily orientated with respect to lightness information and saturation towards hues primarily consisting of green, red, blue or yellow orientated as orthogonal dimensions.

CIE XYZ can be transformed into CIE LUV using the following formula:

Variable_U = (4 * X) / (X + (15 * Y) + (3 * Z))

Variable_V = (9 * Y) / (X + (15 * Y) + (3 * Z))

Variable_Y = Y / 100

If Variable_Y is more than 0.008856 then the following transformation is done:

Variable_Y = Variable_$Y^{(1/3)}$

If Variable_Y is less than 0.008856 then the following transformation is done:

Variable_Y = (7.787 * Variable_Y) + (16 / 116)

Reference_X = 98.07

Reference_Y = 100

Reference_Z = 118.18

Reference_U = (4 * Reference_X) / (Reference_X + (15 * Reference_Y) + (3 * Reference_Z))

Reference_V = (9 * Reference_Y) / (Reference_X + (15 * Reference_Y) + (3 * Reference_Z))

CIE LUV

L = (116 * Variable_Y) – 16

U = 13 * L * (Variable_U – Reference_U)

V = 13 * L * (Variable_V – Reference_V)

This provides Cartesian co-ordinates for CIE LUV space as centred on the reference white point values, providing a separate lightness dimension (L), as well as two opponent colour dimensions relating to green-red saturation (U) and blue-yellow saturation (V). While lightness can range from 0 (Black) to White (100), negative U values relate to green saturation, positive U values relate to red saturation, negative V values relate to blue saturation and positive V values relate to yellow saturation. The maximum values for U and V vary as a function of the lightness value, for instance yellow saturation can reach higher V values in higher lightnesses whereas it is more constrained at lower lightnesses. This space can also be explored using polar co-ordinates as seen in CIE LCh_{uv}, standing for Lightness (0-100), Chroma (0-100) and Hue angle (0-360). Conversions from CIE LUV can be transformed into CIE LCh_{uv} using the following formula, while L remains the same:

$C_{uv} = \sqrt{} (U)^2 + (V)^2$

$h_{uv} = \arctan (V / U)$

The primary advantage of using CIE LUV and LCh_{uv} is the use of a perceptually uniform space, where spatial distances between two colours approximates to their perceptual distance in just-noticeable differences (JNDs) as judged by standard human observer (see fig. 1.11).

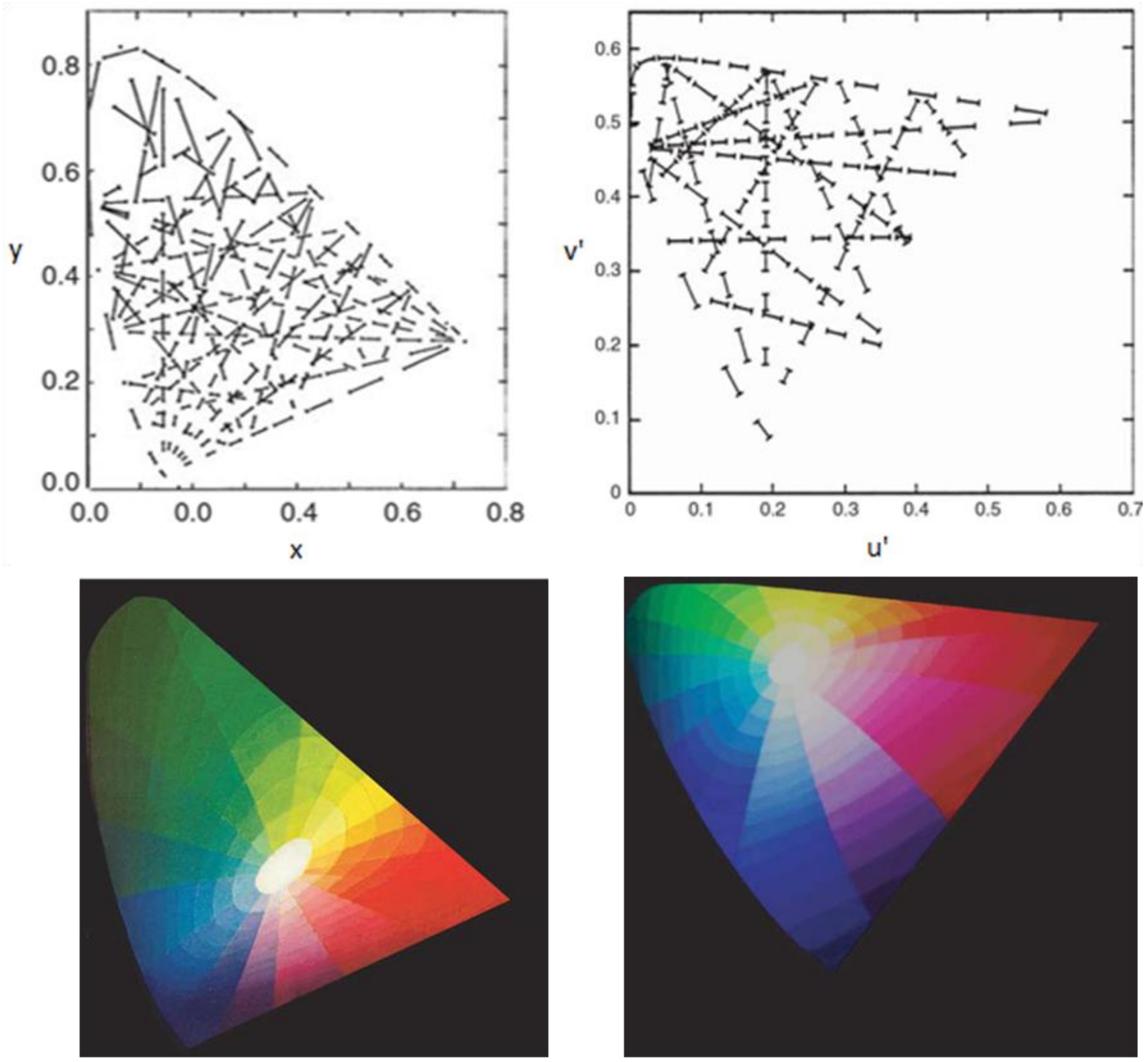


Figure 1.11. Chromaticity diagrams with lines representing 3 just noticeable differences between colour space locations. Upper left chart shows variation in JNDs between colour locations in CIE xy space (Wright, 1941). Upper right chart shows the reduction in variation in JNDs between colour locations in CIE u’ v’ space (also known as the CIE 1960 Universal Chromaticity Diagram, later superseded by the 1976 UCS diagram). Lower pictures are illustrations of the above charts by Louis Condax (Hunt & Pointer, 2011).

The perceptual colour distance between two colours can be worked out in CIE LUV space using the following formula (Robertson, 1977):

Colour difference = $\sqrt{(L_1 - L_2)^2 + (U_1 - U_2)^2 + (V_1 - V_2)^2}$

In chapter 5, participants are given the ability to manipulate the chroma and hue of a single colour in CIE LCh_{uv} space, the effect of the participant's manipulation on the outputted RGB values of the monitor are reached through a reverse of the transformations described above for RGB → XYZ → LUV → LCh_{uv} colour spaces.

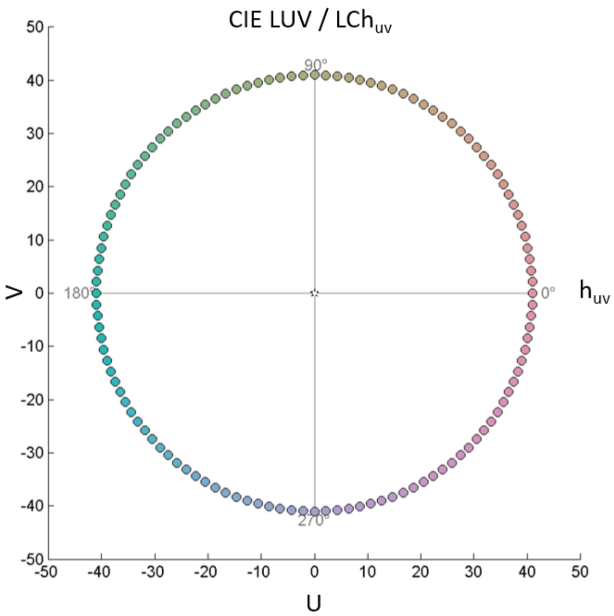


Figure 1.12. Illustration of hue variations within CIE LUV and LCh_{uv} colour space for a single lightness and chroma value.

1.4.11. Alternative colour spaces

There are a multitude of alternative colour spaces that are referenced to a lesser degree in the thesis. This spans alternative device-independent colour spaces to CIE LUV such as CIE LAB, as well as colour spaces commonly used in colour selection contexts such as HSL, HSB and HSV.

1.4.12. CIE LAB

An alternative device-independent colour space to CIE LUV is that of CIE's 1976 CIE LAB which compresses the CIE XYZ colour information using a non-linear cube-root transformation on not only luminance information (like CIE LUV) but the chromatic dimensions as well (unlike CIE LUV). This produces a space with a larger gamut of available colours (even expanding beyond all perceivable colours to hypothetical 'imaginary' colours), and so is suitable for transformations between colour spaces that themselves have different gamuts and would otherwise suffer information loss through a direct transformation. Similar to CIE LUV, this space attempts to be more perceptually uniform than its precursor, CIE XYZ space, so that distances between colour co-ordinates reflect perceived colour distances to the human observer. The three dimensions, L (0-100) represents lightness, whereas the colour dimensions A and B typically span the standard limitation of 256 bytes to span -128 to +127 values. Negative A represents green saturation, positive represents red saturation, negative B represents blue saturation and positive represents yellow saturation, desaturated colours are represented by both A and B being equal to 0. A transformation from CIE XYZ to CIE LAB can be done as follows:

Assuming standard daylight conditions and a 2° Standard Colourimetric Observer, values for illuminant D65 are used as XYZ white point reference values (95.04, 100, 108.88).

Reference_X = 95.04
Reference_Y = 100
Reference_Z = 108.88

Variable_X = X / Reference_X
Variable_Y = Y / Reference_Y
Variable_Z = Z / Reference_Z

If Variable_X is more than 0.008856 then the following transformation is done:
Variable_X = Variable_X$^{(1/3)}$
If Variable_X is less than 0.008856 then the following transformation is done:
Variable_X = (7.787 * Variable_X) + (16 / 116)

If Variable_Y is more than 0.008856 then the following transformation is done:
Variable_Y = Variable_Y$^{(1/3)}$
If Variable_Y is less than 0.008856 then the following transformation is done:
Variable_Y = (7.787 * Variable_Y) + (16 / 116)

If Variable_Z is more than 0.008856 then the following transformation is done:

Variable_Z = Variable_$Z^{(1/3)}$

If Variable_Z is less than 0.008856 then the following transformation is done:

Variable_Z = (7.787 * Variable_Z) + (16 / 116)

CIE LAB

L = (166 * Variable_Y) – 16

A = 500 * (Variable_X – Variable_Y)

B = 200 * (Variable_Y – Variable_Z)

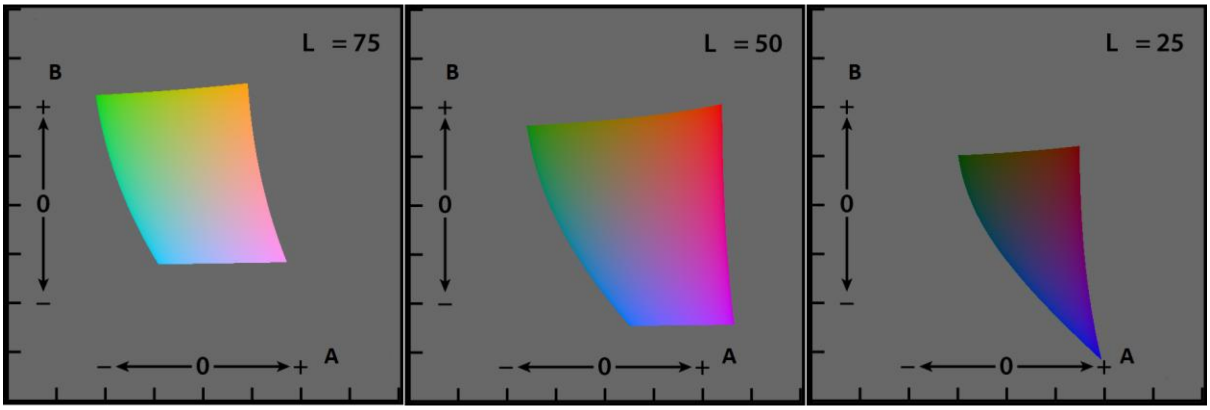


Figure 1.13. CIE LAB colour space at 3 different lightness values. Dimensions denoting A and B range from -128 to +127. Negative A values represent greenness, positive represent redness, while negative B values indicate blueness and positive indicate yellowness. The extent to which the A and B values can be represented is altered as a function of lightness, with more saturated yellows possible at a higher lightness but not at lower lightness (Source: http://en.wikipedia.org/wiki/Lab_color_space, retrieved June 9th 2015).

Perceptual distances between two colours can be worked out using the following formula (Robertson, 1977):

Colour difference = $\sqrt{} (L_1 – L_2)^2 + (A_1 – A_2)^2 + (B_1 – B_2)^2$

1.4.13. HSL, HSB and HSV colour space

Colour spaces used for colour selection applications seek to make movement through colour spaces as intuitive as possible without the consideration of the exact produced colours determined through colour correcting for monitor output. Moving through colours in terms of different hues, saturations and luminances (or brightness / value) using a polar co-ordinate system is immediately practical although likely to contain many minor flaws making it unsuitable for more advanced applications where a device’s outputted colour or colour distances in perceptual space need to be known. HSL space is used in the Creole sensory substitution device in chapter 4, whereas HSV (also known as HSB) space is referenced as belonging to earlier colour SSDs. The transformation from RGB to HSL is described below:

$R_{proportion}$ = R / 255
$G_{proportion}$ = G / 255
$B_{proportion}$ = B / 255

The minimum and maximum R, G or B values are selected
Variable_Min = minimal value of the following ($R_{proportion}$, $G_{proportion}$, $B_{proportion}$)
Variable_Max = maximum value of the following ($R_{proportion}$, $G_{proportion}$, $B_{proportion}$)
Delta_Max = Variable_Max – Variable_Min

L = (Variable_Max + Variable_Min) / 2
If Delta_Max is equal to 0, then saturation and hue also equal 0. If not, the following is continued:

To work out saturation, different formulae are used depending on the level of luminance.
If L < 0.5 then the following is done:
S = (Variable_Max – Variable_Min) / (Variable_Max + Variable_Min)
However, if L > 0.5 then this transformation is done:
S = (Variable_Max – Variable_Min) / (2 – Variable_Max – Variable_Min)

The hue angle is worked out depending on which channel (R, G or B) is equal to Variable_Max.
If $R_{proportion}$ is the same as Variable_Max then:
Variable_H = ($G_{proportion}$ – $B_{proportion}$) / (Variable_Max – Variable_Min)
If $G_{proportion}$ is the same as Variable_Max then:
Variable_H = 2.0 + ($B_{proportion}$ – $R_{proportion}$) / (Variable_Max – Variable_Min)
If $B_{proportion}$ is the same as Variable_Max then:
Variable_H = 4.0 + ($R_{proportion}$ – $G_{proportion}$) / (Variable_Max – Variable_Min)

The Variable_H value is now turned into degrees through the following formula to give the Hue angle in degrees:

H = Variable_H * 60

Alternatives to HSL provide different varieties of luminance information and different interactions with saturation information in providing their final colour. When a colour is completely saturated HSV's value dimension can be at a maximal value (unlike HSL's luminance dimension) and progression to white is defined through a reduction in the saturation dimension rather than the value dimension. As such, high value colours relate to both the bright whites and highly saturated colours (see fig. 1.14).

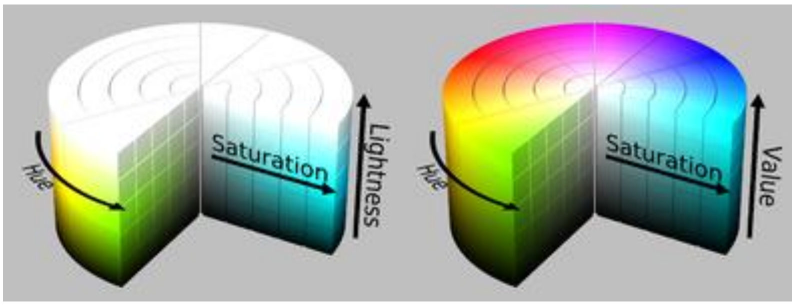


Figure 1.14. HSL and HSV colour spaces. Increasing luminance automatically reduces saturation in HSL, however both saturated colours and bright whites are present as value increases in HSV space (Source: http://en.wikipedia.org/wiki/HSL_and_HSV, retrieved June 9th 2015.

These colour spaces have substantial discrepancies between perceptual judgements made by observers and what the dimension is telling us is equal (Brewer, 1999). As such, while dimensions are named after physical or psychometric qualities, they are instead poor approximations of these and not independent from one another, so only changing the hue angle is likely to also change perceived brightness and saturation of the resulting colours.

2. Representing Colour through Hearing and Touch in Sensory Substitution Devices

2.1. Abstract

Visual sensory substitution devices (SSDs) allow visually-deprived individuals to navigate and recognise the ‘visual world’; SSDs also provide opportunities for psychologists to study modality-independent theories of perception. At present most research has focused on encoding greyscale vision. However at the low spatial resolutions received by SSD users, colour information enhances object-ground segmentation, and provides more stable cues for scene and object recognition. Many attempts have been made to encode colour information in tactile or auditory modalities, but many of these studies exist in isolation. This review brings together a wide variety of tactile and auditory approaches to representing colour. We examine how each device constructs ‘colour’ relative to veridical human colour perception and report previous experiments using these devices. Theoretical approaches to encoding and transferring colour information through sound or touch are discussed for future devices, covering alternative stimulation approaches, perceptually distinct dimensions and intuitive cross-modal correspondences.

2.2. Introduction

2.2.1. What are sensory substitution devices?

Visual sensory substitution devices (SSDs) convey visual information to the senses of hearing or touch with the aim of improving quality-of-life in the blind and visually impaired. At least 57 million people worldwide have a form of visual deprivation that existing medical treatments cannot help (WHO, 2014). Typical SSDs consist of visual sensors such as a web-camera, that relay their information to a coupling device that systematically translates visual dimensions (vertical position, horizontal position, luminance, colour) into tactile or auditory dimensions (frequency, time, intensity) outputted through headphones to the ears or via mechanical / electrical stimulation of the body surface. This is summarised in Fig. 2.1. The user learns to associate different auditory or tactile patterns with objects in their environment. By moving the sensor, the user can also extract information relating to depth and occlusion: for instance, the sensory signal of a nearby object will change more when the device is moved relative to a more distant object. Some of the most commonly studied SSDs are the 'vOICe' (Meijer, 1992) and BrainPort (Bach-y-Rita, Kaczmarek, Tyler & Garcia-Lara, 1998). In the vOICe, each pixel in the image is represented by a sound. The vertical position of a pixel is translated into frequency and the luminance of the pixel is translated into amplitude. The horizontal position is translated into both time and stereo-panning such that the image is heard piecemeal, typically over 1 second. The sound of each scan-through is referred to as a 'soundscape.' In the BrainPort, there is a one-to-one spatial correspondence between pixels in the image and location in an array of electrical stimulators (applied to the tongue) with luminance translated into the intensity of stimulation. In these devices there is no translation of colour information into hearing or touch, beyond the vOICe's option to translate a central colour into a spoken colour name. However, a variety of other devices have been developed that do not discard colour information. These devices are reviewed here. In addition, we consider various options for how colour could be represented in future devices with particular reference to using perceptually distinct dimensions and the psychological literature on intuitive cross-modal mappings between colour and sound / touch.

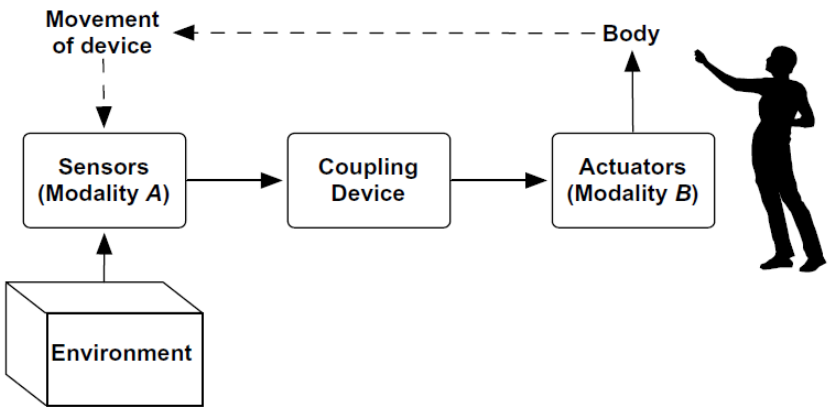


Figure 2.1. Structure of a sensory substitution system (from Visell, 2009).

2.2.2. The utility of colour information in SSDs

There are several reasons why the blind and visually impaired may benefit from the presence of colour in SSDs. Firstly, colour information provides clues to object identity (e.g. bananas are yellow). Secondly, access to colour information enables communication between the blind and sighted about the visual world. Finally, colour information is important for figure-ground segmentation (Goffaux et al., 2005; Rousselet, Joubert & Fabre-Thorpe, 2005; Torralba, 2009). Rivest and Cavanagh (1996) investigated the contributions of visual dimensions (luminance, colour, texture, motion) to defining object boundaries. They found that no visual dimension played a privileged role in localising object contours but note that "discontinuities in luminance created by shadows are not reliably linked to object contours, whereas continuities in other attributes (e.g. colour, motion and texture) are much more reliably linked to object contours." Crucially, the utility of colour information for visual scene/object recognition is greater for medium and low resolutions than it is for higher resolutions (Torralba, 2009) – see Fig. 2.2. Most sensory substitution devices have a relatively low resolution being constrained by both technological limitations and the users' perceptual ability. The upper-bound resolution of tactile devices is determined by the number of stimulators whereas in auditory devices it depends on the software itself (the algorithm for down-sampling an image into pixels). In reality, the resolution is determined by limitations of the users' perceptual system and is typically lower. Snellen visual acuity tests for blind SSD users have shown that users can reach 20/860 (20/430 with training) and 20/200 on 'E'-orientation tests using tactile and audio SSDs respectively (Chebat, Rainville, Kupers & Ptito, 2007; Haigh, Brown, Meijer & Proulx, 2013; Sampaio et al., 2001; Striem-Amit, Guendelman & Amedi, 2012).

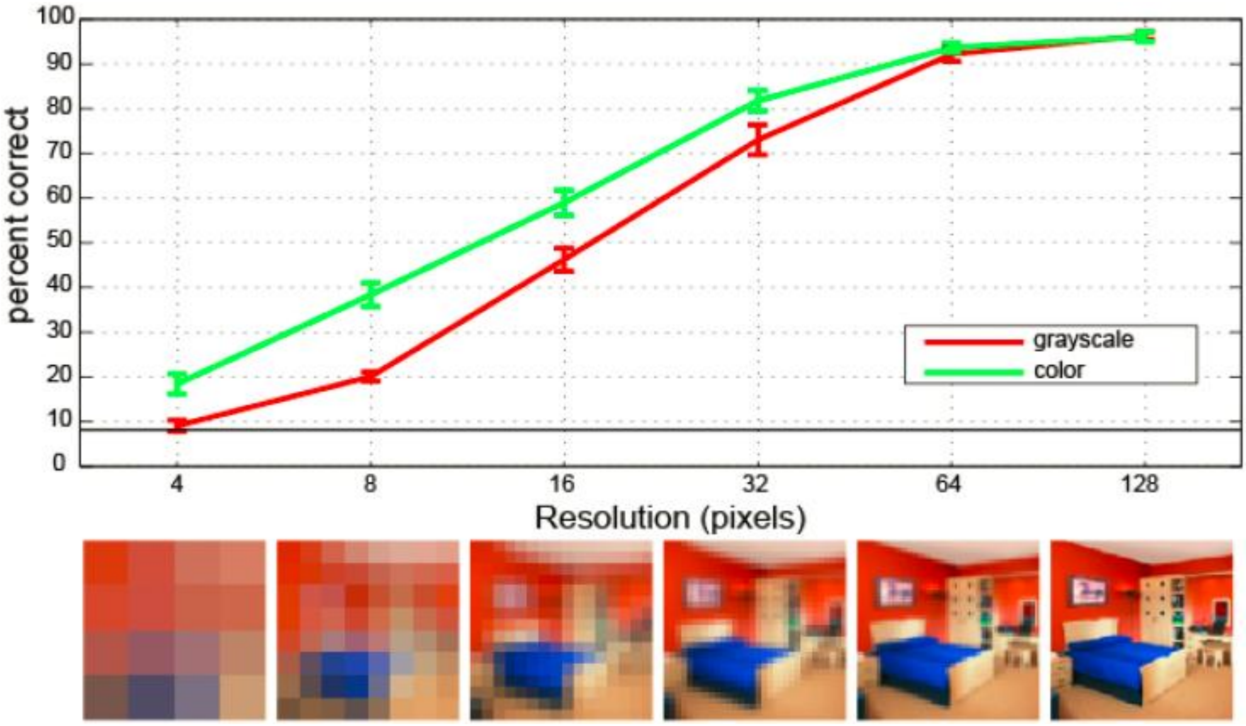


Figure 2.2. Correct scene identification from 240 images split into 12 categories. Bottom line indicates chance performance (8.3%), a minority of images are ambiguous in category, making 100% correct identification unlikely. Error bars indicate 1SE. (from Torralba, 2009)

The most common computational representation of colour is RGB-space (adding mixtures of Red, Green and Blue create individual colours). This can be used as the basis for colour representation in SSD systems, or it can be converted to other colour-spaces. Some devices represent colour as a linear variable (emulating the red to violet electromagnetic frequency), others use discrete symbolic categories (e.g. language or codes), and some create intuitive variables such as hue, saturation and luminance (HSL or HSB-space) or attempt to mimic human colour space (CIE LAB, CIE LUV). Human colour perception has a structure that differs from both its linguistic representation and the physical properties of colour (Bornstein, 1987). Veridical colour perception has several distinctive features:

- Focal colours, which are the best exemplar of a colour category (Regier, Kay & Cook, 2005).
- Categorical perception, where within-category variation is minimised while between-category variation is exaggerated (Harnad, 1987).
- Red-green and blue-yellow colour opponency, where excitation of one colour, inhibits its opponent (Engel, 1999; Hurvich & Jameson, 1955, 1957).
- Illuminating context, such that colour perception takes into account environmental luminance (Hansen, Walker & Gegenfurtner, 2007).
- A circular representation of hue such that violet and red are perceptually close rather than distant.

As such, deciding which structure of colour to replicate has important implications for both its encoding but also its application.

The different uses of colour described above would necessitate different spatial resolutions (single or multiple pixels being represented) and different kinds of colour representation (noted here as symbolic, linear or circular). Devices that only encode a single point of colour would aid in recognition of known objects such as a yellow banana or facilitate communication with a sighted person on clothing. However, figure-ground segmentation requires multiple colours to be translated, presented either simultaneously or mapped out over time so users can detect boundaries through contrast. Current SSDs differ in this regard as well as with respect to the system of colour coding. A symbolic representation of colour (e.g. discrete categories of 'yellow,' 'brown,' etc.) is sufficient for figure-ground segmentation. However, within-category variations would be undetectable (e.g. different reds under the same label) and there is no necessary ordering of colours as the labels are arbitrary markers. As a result, yellow and orange would not be recognised as being perceptually similar. Symbolic coding is used in devices such as the 'Haptic-Colour Glove' SSD (Kahol, French, Bratton & Panchanathan, 2006), whereby distinct vibrotactile patterns similar to Morse code indicate a colour category. In some systems, hue is represented as a linear variable (from red to violet) in analogy to their physical electromagnetic frequency. This allows a representation on a single variable (e.g. frequency) whilst retaining the ability to detect multiple hues and their ordering. Linear coding occurs in the 'EyeBorg' SSD (Else, 2012), with increasing auditory frequency communicating increases on the hue variable to the user. In other systems, hue is represented as a circular variable, where a smooth violet-red transition is kept and the perceptual distance between colours can be assessed, similar to veridical colour perception. The 'See ColOr' SSD (Bologna et al., 2007) provides a cyclical representation through utilising unique sounds for each focal colour and indicating transitions between neighbouring colours through combining sounds together. Systems can also differ in terms of whether saturation and luminance are represented in addition to hue. There have been many approaches to representing colour through touch and hearing and these devices are summarised in tables 2.1 and 2.2. There is not necessarily a 'correct' representation; it depends on the function it is designed for and it is, to a large extent, an empirically open question as to which is most useful.

One source of information is how navigation works in naturalistic environments with varying luminance, saturation and hue. Luminance and hue information have been found to be reliable for computer navigation and recognition in shadowed environments (Crisman & Thorpe, 1993; Orwell, Remagnino & Jones, 2001). However greyscale SSDs have only been tested in artificial high-contrast greyscale environments with consistent lighting conditions (Chebat et al., 2011; Segond, Weiss &

Sampaio, 2005), and therefore their effectiveness in navigating real-world environments is unproven. The addition of cues more resistant to changes in natural illumination such as hue, may allow more stable object identifiers for the user. This reduces the need for information on light sources, their location and interaction with the environment, an ability which is likely impaired with low spatial resolution 'vision.'

2.3. Representing colour in vision-to-tactile sensory substitution devices

We first consider existing devices and solutions before turning to consider potential directions for the future.

2.3.1. Existing devices

Some existing solutions for representing colour via touch / haptics include: translating colour into texture (smoothness, degree of resistance): translating different dimensions of colour space (e.g. R, G and B channels of a single pixel) into spatially distinct tactors; or translating colour (as a category or linear dimension) into different kinds or levels of stimulation (frequency, intensity) on a single tactor. The devices are described in more detail below and summarised in Table 2.1.

2.3.1.1. TACTile Image Creation System or TACTICS (1997)

Way and Barner's (1997b) 'TACTile image creation system' prints out 2D images into an embossed representation on paper for tactile exploration. Luminance-defined edges appear as embossed lines. Different colours were proposed to be translated into different textures; however, this was never actualised and instead provided inspiration for several later devices. Although this device has limited utility in terms of real time recognition due to requiring a printed output, the concept could be developed using tablets with real-time tactile feedback (Roth, Richoz, Petrucci & Pun, 2001). The translation of colour could be done through 'artificial textures' that preferentially stimulate specific tactile receptors (Kajimoto, Kawakami & Tachi, 2004), temperature (Hribar & Pawluk, 2011) or sound (Bologna et al., 2007; Ebert, 2005).

2.3.1.2. Phantasticon (1999) and PHANToM Texture-Colour SSD (2006)

Sjöström and Rassmus-Gröhn (1999) reported a computer-haptic interface using PHANToM - a joystick that moves in a 3D space or 'haptic scene'. Different colours are represented by different levels of resistance as the device is moved thereby creating a sense of texture. As with TACTICS, the specific texture-colour mappings are not reported.

One later device that also uses the Phantom joystick divides colour into four dimensions (R, G, B and luminance) which can be explored haptically one at a time (Kahol et al., 2006; Fig. 2.3). Increased resistance corresponds to increased presence of that dimension and the haptic dimensions are limited to off, medium and high levels of friction. So when exploring an image in 'R' mode, the presence of a red patch would feel like high resistance (sticky), orange would have a medium resistance and black would have low resistance (smooth). Initial experiments featured three groups of five trained participants; blind, sighted (who also did it blindfolded afterwards), and always blindfolded. Participants were presented with one of 21 different colours (ROYGBV and grey at three luminance levels each). Recognition of individual colours on the first trial was near 100% accuracy for sighted participants and lowest at 90% accuracy for initially blindfolded participants. By trial 4, all groups had reached near 100% accuracy. Perceptual similarity between colours was assessed through previous participants with 15 new untrained participants (who are unaware colour is being represented) rating the similarity of two colours through the device. This was mapped using multi-dimensional scaling to plot a similarity matrix that groups perceptually similar colours together which was then compared to HSV colour space's structure. They reported a high level of perceptual similarity between all groups and HSV space with no significant difference between trained and naïve groups, which for the latter group suggests a retention of perceptual space independent of colour knowledge.

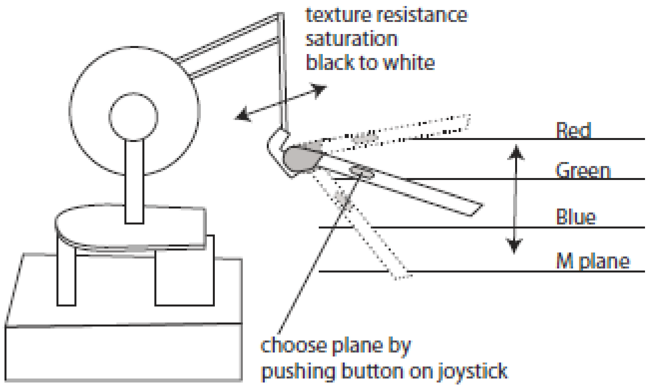


Figure 2.3. PHANToM texture-colour SSD (Kahol et al., 2006). A single point of colour on the PC is represented in RGB space and converted to separate red, green, blue and luminance (M plane) dimensions for interrogation by the user. Planes are selected individually via a push-button, and lateral resistance to movement indicates the presence of a low value (free movement) or high value (resistance to movement) on that colour dimension.

2.3.1.3. VIDET Glove (1998) and Haptic-Colour Glove (2006)

Cappelletti, Ferri and Nicoletti (1998) represented a single point in RGB colour space on three fingertips (relating to red, green and blue dimensions) as well as by varying vibrotactile frequency (low, medium, high amplitude). So the presence of red would be felt as high vibration intensity on the 'red finger', yellow would be felt as high vibration intensity on the red and green fingers (because yellow is represented by red and green together in RGB space), and so on. Blind participants were tasked with identifying colours after training (using passive tactile stimulation) and identifying one of three coloured geometric shapes using a graphics tablet. Performance ranged from 81-100% for colour recognition and was 100% for shape recognition. Participants were noted to make errors that confused visually similar colours (e.g. mistaking an orange for a yellow or a red). Habituation to vibrations also increased response times, indicating that continual use impairs the tactile response, decreasing the user's resolution.

Kahol et al. (2006)'s glove system (Fig. 2.4) is very similar but uses additional sites to represent high/low luminance (thumb / little finger respectively) in addition to R, G and B being felt on the index, middle and ring fingers. The involvement of each dimension is represented through four levels of temporal vibration patterns (similar to Morse code) presented in parallel. Kahol et al. (2006) report that five trained blind users correctly identified 21 colours with near perfect accuracy and, while exploring colour-simplified natural scenes, participants maintained this near-perfect accuracy for colour-naming. Further tests on 15 sighted participants compared the RGB output strategies of the Phantom's serial approach to the glove's parallel approach. They found that learning specific colours is more difficult for the glove's parallel presentation (78.5% accuracy) than the Phantom's serial approach (100% accuracy).

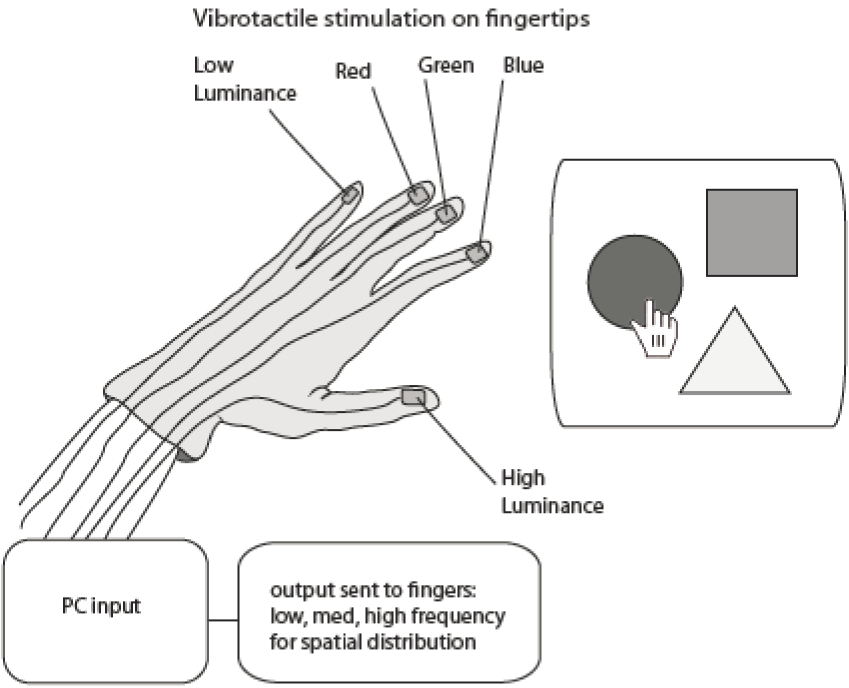


Figure 2.4. Haptic-Colour Glove (Kahol et al., 2006). Colours selected through a computer program are outputted through to vibrations on different fingers of the VIDET glove, the spatial location of the vibrations indicating the presence of red, green, blue and the brightness of the selected colour.

2.3.1.4. Chromo-Haptic Sensor-Tactor Device, CHST (2008)

The Chromo-Haptic Sensor-Tactor (CHST; Fig. 2.5) device has four short-range fingertip-mounted colour sensors as part of a glove (Tapson et al., 2008). The four finger-mounted sensors are tuned to detect four dimensions each: R, G, B and luminance. Depending on the translation method used, the CHST can represent between one to four colours simultaneously. The information is relayed to four belt-mounted vibrotactile stimulators (T1-4), varying in vibration and temporal modulation to convey colour information. Unlike the initial attempts to transfer colour information, this device is the first to use a colour sensor for the external environment rather than a pc camera. The CHST uses HSV space (Hue, Saturation, Value), with hue deriving four colour categories (red, green, blue, yellow). These colour categories are encoded in a variety of symbolic representations that vary dependent on the number of simultaneous colours conveyed to the user. A single colour is conveyed via colour-tactor pairings that vary spatially on the skin (i.e. yellow = T1, red = T2); four colours use sensor-tactor pairings, with colour indicated via temporal patterns (i.e. sensor 1 on T1, sensor 2 on T2, with blue = pulse, pulse). Finally, conveying two colours simultaneously use both spatial and temporal changes to convey colour (i.e. sensor 1 uses T1 and T2, sensor 2 uses T3 and T4, with blue = alternating tactor vibrations). All methods involved symbolic codes rather than replicating any colour space structure. Participants were evaluated on learning red, green, blue and yellow with all these coding schemes, with all methods having participants able to reach 100% colour-identification accuracy. However the spatial method was fastest.

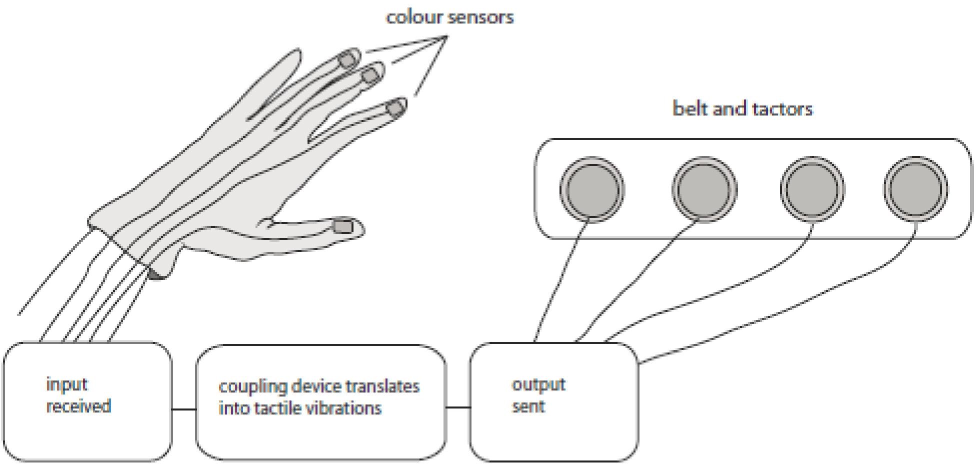


Figure 2.5. Chromo-Haptic Sensor-Tactor (Tapson et al., 2008). An external colour sensor records colour in RGB space with luminance, which is then converted to categories of common colours, then provided as a symbolic code using the belt-mounted tactors.

Subsequent models (Schwerdt, Tapson & Etienne-Cummings, 2009; Fig. 2.6) replaced symbolic coding with hue and saturation (or brightness) values being logarithmically mapped to vibrotactile-frequency (0-33Hz) on separate arm-mounted tactors. By using HSB colour space and a continuous dimension for output, they hoped to improve the structure, intuitiveness and resolution of colour representation. However since frequency is a linear value and hue is cyclical, this creates a break between colour neighbours (i.e. HSB neighbouring colours violet and red represented furthest apart in frequencies).

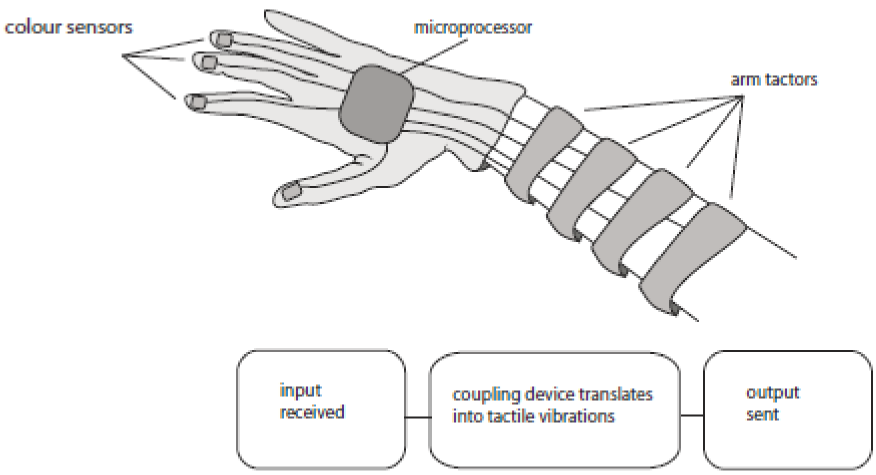


Figure 2.6. Chromo-Haptic Sensor-Tactor (Schwerdt et al., 2009). This version of the device converts received RGB colours from the external sensors to be represented in HSB colour space using frequency vibrations on the arm-mounted tactors.

2.3.1.5. Finger-mounted colour SSD (2012)

Burch (2012; Burch & Pawluk, 2009) describes a device with multiple finger-mounted RGB sensors that compares RGB values with pre-set colour categories which have associated vibrotactile stimulations delivered to the fingertip (Fig. 2.7). The encoding of colour is achieved through varying a square wave vibrational signal, with one of three temporal frequencies (30, 60 and 120Hz), each of which has three duty cycles (30, 50 and 70% 'on' during each cycle) creating nine distinguishable colour categories. Moving the device across various colour contrasts results in the device achieving a 1 to 2mm spatial resolution for participants, with lower saturation-contrasts resulting in slightly reduced discrimination. Burch describes an experiment where 19 blind participants explored 12 abstract images, each containing 3-5 separate sections denoted by texture orientation (horizontal, vertical or diagonal coloured lines) or single colour patches. Participants were given a free choice on the number of distinctive sections and their orientation, on average reaching above 90% accuracy for both the number of sections and their texture. Whilst colour identification was not evaluated, the ability to discriminate accurately using motor movements and frequency / duty cycles marks these dimensions viable for future tactile SSD use.

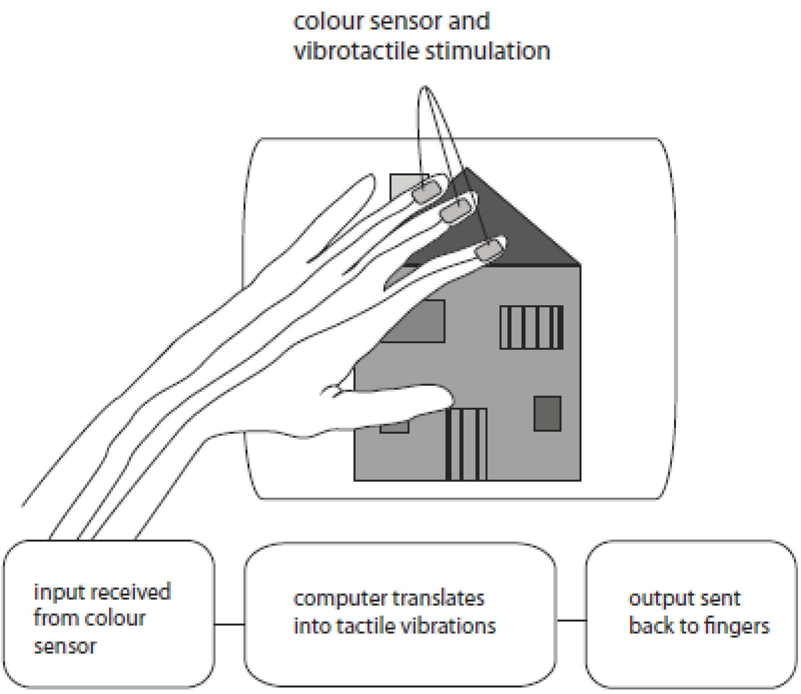


Figure 2.7. Finger-mounted colour SSD (Burch, 2012). Different shades of grey represent different colours on an image. These colours are sensed by the finger-mounted sensor and after processing are outputted as vibrational patterns on the same finger as the original sensor.

2.3.1.6. Electro-Neural Vision System, ENVS (2004)

The ENVS divides visual space into ten sections, each conveying colour and depth information through electro-tactile pulses and intensity to the fingers (Meers & Ward 2004, 2005 &

2007; Fig. 2.8). Depth information is created through a disparity map between head-mounted stereo cameras, which is then conveyed through the intensity of electro-tactile stimulation (between 30 to 80V). Pulse frequency is determined by the average colour within that section of space. The average colour was originally mapped between its location on the electromagnetic spectrum and frequencies of between 10 and 120Hz. Initial demonstrations showed that participants struggled to detect both kinds of information (intensity and frequency). As such, the final version moved to a system of colour categories linked to discrete frequencies. The ten (adjustable) horizontal sections of space (and related depth / colour information) are divided amongst the fingers, so that holding the fingers out in front relates to the spatial position of information beyond peripersonal space. While in its default state only the horizontal axis is constantly updated, head movements can provide the vertical axis. Informal navigation experiments were reported by Meers and Ward (2005) with trained blindfolded users in a familiar environment. Participants were able to navigate several corridors and identify an appropriate target (a blue door) located between a grey filing cabinet and red fire extinguisher.

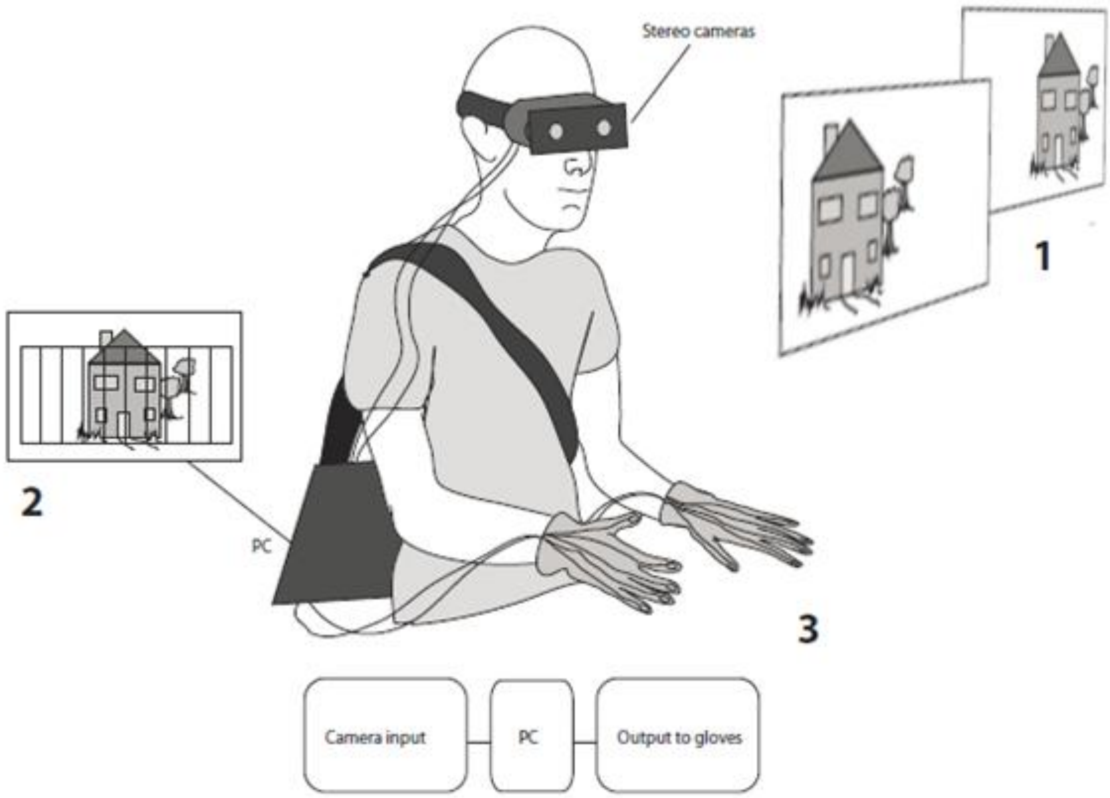


Figure 2.8. Electro-Neural Vision System (Meers & Ward, 2004). Head-mounted stereo-cameras pick up two 2D images (1) that form a disparity map to infer depth information and a final 2D image (2) from which the dominant colours in 10 sections of space are encoded as electrical stimulation for 10 fingers in the ENVS-gloves output (3).

Table 2.1. Coding strategies in current tactile-colour SSDs.

SSD	Spatial Representation	Colour Representation	Tactile-Colour Coding	Reference
TACTile ICS	Fingertips over 2D embossed paper printout.	Luminance informs edge detection. Colour discussed but not implemented.	Texture mapping proposed[2].	Way & Barner, 1997b.
Phantasticon	Single point, PC input.	Specifics not discussed.	Lateral joystick resistance.	Sjöström & Rassmus-Gröhn, 1998.
PHANToM T-V SSD	Single point, PC input.	Symbolic - Red, green, blue and brightness dimensions converted to distinct categories, presented sequentially.	Lateral joystick resistance.	Kahol et al., 2006.
VIDET Glove	Single point, PC input.	Symbolic – Red, green and blue dimensions converted into distinct categories, presented simultaneously.	Amplitude[3] and skin location.	Cappelletti et al. 1998.
Haptic-Colour Glove	Single point, PC input.	Symbolic – Red, green, blue and luminance dimensions converted into distinct categories, presented simultaneously.	Pulse sequence and skin location.	Kahol et al., 2006.
Chromo-Haptic Sensor-Tactor	Multiple external sensors, which can be horizontally or vertically aligned.	1st version is symbolic, converting hue, saturation and value into distinct categories. 2nd version is linear and symbolic, using hue and saturation as linear variables and converting brightness into categories.	Pulse sequence and/or skin location.	Tapson et al., 2008. Schwerdt et al., 2009.
Finger-Mounted Colour SSD	Multiple finger-mounted sensors, which can be horizontally or vertically aligned.	Symbolic – Red, green, blue and luminance dimensions converted into distinct categories.	Vibrotactile frequency[1,3] and duty cycles.	Burch & Pawluk, 2009. Burch, 2012.
ENVS	Ten points of space that can be vertically or horizontally aligned.	1st version is linear, simplifying colour values into a 'colour spectrum.' 2nd version is symbolic, using red, green, blue and luminance dimensions to derive categories.	Pulse frequency and skin location.	Meers & Ward, 2004, 2005 & 2007.

Supported as a correspondence by: [1]Martino & Marks, 2000; [2]Ludwig & Simner, 2013; [3]Ward, Banissy & Jonas, 2008.

2.3.2. Potential future directions

In this section we consider different ways in which colour could be conveyed by touch and haptics that have not yet been developed. This includes the role of psychology in assisting the design of such devices, representing hue in terms of circularly varying dimensions of touch, stimulating different kinds of skin receptor to represent different colours, using temperature to represent colour and employing haptic feedback on touch-screen technology.

The development of these technologies can be informed by research in psychology concerning cross-modal correspondences between colour and touch. Pre-existing intuitive associations between modalities (whether culturally specific or not) may help to bridge the modality-gap. For instance, there are established associations between colour and temperature as well as roundness, smoothness, and softness dimensions (Ludwig & Simner, 2013; Morgan, Goodson & Jones, 1975). Based on the findings of Ludwig and Simner (2013), one could envision luminant elements of a visual image represented through smooth and soft textures on a tactile display. While for children and adolescents correspondences for saturated colours were found this was not the case for adults, leaving the door open for alternative correspondences for these colour dimensions. It could be the case that the explicit use of correspondences to represent visual characteristics might maintain soft-chroma and smooth-chroma correspondences into adulthood as would be predicted by the statistical explanation of correspondences (Spence, 2011). Since these kinds of associations tend not to be purely linguistic but also affect perceptual judgements (Martino & Marks, 2000) if visual judgements (e.g. 'red') are congruent with their intuitive direction from cross-modal correspondences (e.g. hot) to what extent would this aid learning, processing time and the phenomenological experience relative to avoiding the use of correspondences? For beginners of such devices, would the use of correspondences improve the aesthetics of using such a device, as well as sharing in additional common understandings of colour with the sighted such as emotional valence (Walker et al., 2013)? For individuals with no experience of sight, is the use of correspondences to communicate as many aspects of colour desirable or is the information required to discriminate two different colours enough? If correspondences can reduce the cognitive effort required to learn and master these mappings, to what extent would any effects of this be modulated by prior visual experience? The answers to these psychological questions may play a role in fostering long term adoption of these devices outside of the lab environment creating opportunities to research the long term impact of sensory tools on perceptual and neurological development.

Evidence of cross-modal associations present in both acquired and developmental synaesthesia can further illustrate optimal SSD design both through making implicit correspondences explicit and easier to identify (Ludwig & Simner, 2013; Simner & Ludwig, 2012; Ward et al., 2006), but also through illustrating potential 'routes of least resistance' in the brain to forming stable long-term cross-modal connections through the patterns (both sensory and neurological) present across groups of synaesthetes (Afra et al., 2009; Armel & Ramachandran, 1999; Kupers et al., 2006; Merabet et al., 2009; Ro et al., 2007; Rouw et al., 2011). The use of and creation of these routes might be easier to produce long-lasting changes in the perceptual processing of SSD experts. For instance, the prevalence of both acquired and developmental touch-colour synaesthesia relative to sound-colour synaesthesia may parallel the difficulty in producing 'visual' sensations in SSD experts that use touch or sound to communicate visual information (Ortiz et al., 2011; Ward & Meijer, 2010).

Another way that psychological research can assist in SSD design is by specifying the most perceptually distinct dimensions with which to work and how perceptual resolution may change along this representation. For instance, one possibility for a circular representation of colour on a single tactor is to code colour in terms of two linear dimensions that interact (e.g. vibrotactile frequency and intensity). Whilst some existing devices have utilised frequency and intensity, they have not incorporated hue as a circular variable. This could be achieved by creating a 2 dimensional space defined by frequency (low-medium-high) and intensity (low-medium-high) but such that opponent colours are aligned with the two axes (e.g. frequency with green-red, amplitude with blue-yellow). This creates focal points for perceptually distinct tactile dimensions to line up with the purest exemplar of a colour (e.g. highest frequency = red, lowest = green). This is illustrated in Fig. 2.9. Whether this will be useful depends on several factors such as on the perceptual ability to discriminate independently frequency and intensity. Whilst perceivers find this difficult (Way & Barner, 1997a), it should be possible to calibrate the device such that the intervals are determined psychophysically rather than physically (Morley & Rowe, 1990; Roy & Hollins, 1998; Taylor, 1977; Verrillo, Fraioli & Smith, 1969) thereby minimalising problems such as adaptation to vibrotactile stimulation (Hollins, Goble, Whitsel & Tommerdahl, 1990). Another potential issue is the extent to which this kind of coding is intuitive, given the perceptual tendency to link vibrotactile frequency to luminance rather than hue in sighted individuals (Martino & Marks, 2000).

Another consideration is the interaction between the device and point of contact. The skin contains four types of tactile mechanoreceptor characterised by their skin location, adaptation rate, receptive fields and typical usage (Vallbo & Johansson, 1984). Kajimoto et al. (2004) describe the process of using electrotactile stimulation to produce sensations of 'natural' tactile-stimulation through stimulating each mechanoreceptor separately. This has the potential to encode different dimensions of colour space (e.g. R, G and B in RGB space) by stimulating different kinds of mechanoreceptor. Perceptually this creates discrete categories of experience such as hard pressure, soft pressure, hard pressure + vibrations, soft pressure + vibrations, low vibrations and high vibrations. In combination with temporal and spatial information from the skin, multiple dimensions of visual content (including but not limited to colour) could be encoded.

Using different temperatures to represent colour is presently an unexplored approach, even though there are, at least in Western culture, associations between colour and temperature (Morgan et al., 1975). To give an example of a potential device, Yang, Kyung, Srinivasan and Kwon (2006) describe a five fingered device that delivers either vibrotactile feedback or thermal feedback using Peltier thermoelectric heat pumps. It maintains a temporal resolution of 6°C/s cooling rate and 12°C/s heating rate, with a range between 0-60°C. While this approach has been used to simulate thermal diffusivity of different materials, it could be used for the representation of visual information such as 'warm' or 'cool' colours (Ou et al., 2010).

Finally, the popularity of touch-screen technology in phones, tablets and computers provides an accessible way to examine precise spatial co-ordinates with tactile feedback to the user. Despite the finger being one of the most spatially sensitive regions of the human body with precise motor control, its use in SSDs have been limited due to its relatively small surface area and regular use in daily activities. However the use of a single fingertip on touch-screen technology allows the user to examine digital visual images in a progressive manner. Through examining the tactile-feedback from a series of smaller points the user can explore and construct the larger image over time. This allows the user to focus on points of interest or difficulty, much in the way eye movements reflect our visual attention. Furthermore, multi-modal integration between finger explorations, tactile feedback and auditory information can allow multiple modalities to undertake tasks that they are best suited for, whether through discrimination, bandwidth or intuitive associations.

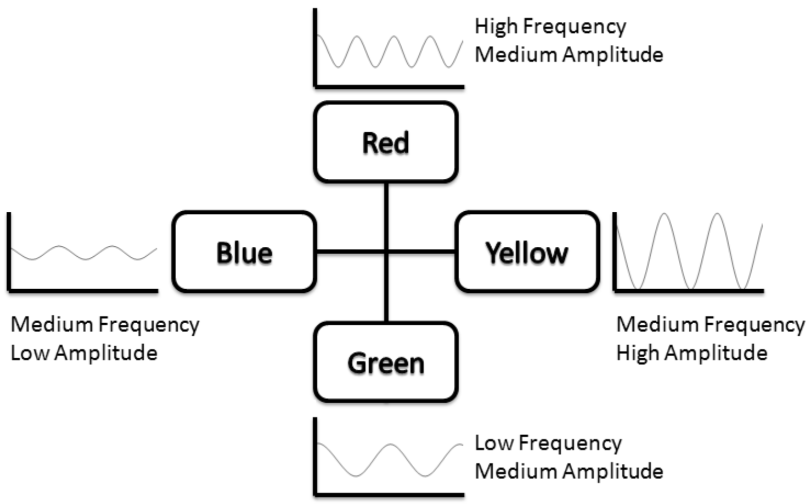


Figure 2.9. Vibro/electro-tactile dimensions of frequency and amplitude provide focal points of perceptual distinctiveness that mimic the hue-circle based on R-G and B-Y opponencies. The X-Axis shows amplitude as the most distinctive dimension, while the Y-Axis shows frequency as the most distinctive dimension.

2.4. Representing colour in vision-to-auditory sensory substitution devices

As with the touch-based devices described previously we will first consider existing solutions to colour-sound conversion in SSDs and then consider potential future directions.

2.4.1. Existing devices

Some devices convert a single (or averaged) point in space into spoken colour names (McMorrow, Wang & Whelan, 1997). Although of potential interest and value, this does not necessarily meet some definitions of a sensory substitution device. Ward and Wright (2014) argue that such devices should operate by translating between senses at the perceptual rather than symbolic level. For instance, many of the devices described below convert some aspect of colour to auditory dimensions such as pitch, loudness and timbre. Timbre itself refers to the quality of a complex sound that allows different instruments or vocals to be discriminated even if they have the same pitch and loudness. The devices are described in more detail below and summarised in Table 2.2.

2.4.1.1. Soundview (2003)

The Soundview has participants select a single pixel at a time of a hidden image on a tablet using a stylus. The colour content of this pixel is represented as subtle variations of noise which is then scaled by the velocity of haptic exploration by the user mimicking the sound of scraping a surface (Doel, 2003). HSB colour space is used and represented through applying various filtering techniques to white noise (a combination of all audible frequencies). First, brightness is represented by altering a lowpass filtering cutoff frequency for the white noise with bright colours retaining most frequencies while dark colours retain only lower frequencies. If there is saturation Reson filters are applied. These filters emphasise a specific frequency and reduce the amplitude of all other frequencies. The level of saturation determines the extent to which a single frequency is emphasised, so that de-saturated colours have all remaining frequencies loud, while saturated colours have a single prominent frequency. The hue of a colour selects which frequency the Reson filter emphasises. A distribution of 12 frequencies that span a single octave are chosen, and a Reson filter can be used to emphasise one of these. By shifting emphasis between the highest frequency Reson filter and lowest frequency Reson filter, a perceptually smooth transition is made between these distant frequencies. This occurs due to the ‘Shepard illusion,' whereby smooth progressions between the highest pitch note in an octave to the lowest pitched note in the same octave produces a sensation of continually rising pitch, this allows a perceptually smooth transition between the highest and lowest frequencies in order to create a perceptually cyclical representation (Shepard, 1964). Further work with the Soundview has identified its fast learning rate for 2D shape identification in comparison to the greyscale vOICe image sonification system (Doel et al., 2004).

2.4.1.2. Eyeborg (2004)

The Eyeborg system was developed to represent colours from a head-mounted webcam as sounds for a single case, NH, who has congenital achromatopsia. Since 2004, NH has worn several versions of the Eyeborg device. The original algorithm mapped a linear (red-to-violet) representation of hue to 25 musical tones (2004 version). From 2007, Eyeborg outputted 360 micro-tones spanning an octave (reflecting degrees of the Hue wheel in HSB space) with additional saturation-loudness mapping. Finally, in 2010, a self-contained Eyeborg chip was created, which now sensed infrared and ultraviolet frequencies and a new output was chosen through occipital bone conduction. Information from NH's greyscale vision and Eyeborg's vibrations can be integrated by linking the spatial location of the seen object with the spatial position of the webcam's focus; however, colour is not reported to invade his visual space. NH describes his SSD expertise as "completely normal to have this constant sound around me" (Else, 2012). His reported phenomenological experience is that of "hearing colour" rather than seeing it, with perceptually similar sounds to Eyeborg being automatically interpreted as Eyeborg signals. As such, the sound from an object and the sound of an object's colour can create a source of auditory interference. For instance a blue-coloured vacuum cleaner may make a sound linked to red (NH, personal communication, August 29th, 2012).

2.4.1.3. Kromophone (2009)

Kromophone simultaneously sonifies the individual components (RGBY and greyscale) that make up a single point of colour (Capalbo & Glenney, 2009). In the final version of the device, each focal colour (red, green, blue, yellow, white, grey, black) has its own characteristic sound comprising of pitch, timbre and panning (e.g. 'red' is a high pitched trumpet sound emanating from the right of space, grey would be a middle pitched sine wave in stereo). Each 'colour' is a sum of the focal sounds with the intensity (i.e. loudness) of each component determined by its contribution to the final colour. So an orange colour would be a mixture of red and yellow sounds, each of moderate loudness. To aide in navigation, horizontal sweeps (alongside user-controlled area-size and light levels) were added to create a horizontal axis over time. While the original version of the Kromophone sonified HSL colour space (mapping hue to pitch, saturation to left-right panning and luminance to volume), this was discarded after it was discovered only participants with prior musical training were proficient with this coding.

Capalbo and Glenney's experiments with the kromophone suggest that it outperforms the vOICe in search, discrimination and navigation tasks for blindfolded sighted participants. Luminance localization in a dark room had equal response times between the Kromophone and vOICe, however even slight increases in environmental luminance resulted in a sharp drop-off in performance for vOICe users only.

2.4.1.4. Seeing Colours with an Orchestra, See ColOr (2007)

The See ColOr device (Bologna et al., 2007) extracts a 25 pixel horizontal line from a videocamera and simultaneously sonifies each pixel's HSL values using the following rules:

- Hue is specified using a unique timbre for each colour: Red (Oboe), Orange (Viola), Yellow (Pizzicato), Green (Flute), Cyan (Trumpet), Blue (Piano) and Purple (Saxophone). Between-category hues mix the volume of the neighbouring hue-timbres, allowing for smooth transitions.
- Saturation is denoted through the pitch of the hue's instrument: 0-24% (C note), 25-49% (G note), 50-74% (B flat note) and 75-100% saturation (E note).
- Luminance is denoted through the addition of either a double bass (<50% luminance) or a singing voice (>50% luminance) to the sound mix, which are sounded in one of the following notes in order of increasing brightness; C (darkest variant), G, B-flat or E (brightest variant). In addition, luminance values near 0% discard all other sounds in favour of double bass. Near 100% luminance sonifies only the singing voice but, at 50% luminance both double bass and singing voice are discarded, allowing only hue's timbre (and saturation-note) to be played.

Communicating the spatial distribution of the 25 sonified pixels is denoted through inter-aural time delays as well as the intensity differences between the ears that are used in natural sound localization (Wang & Brown, 2006). Depth perception was later incorporated through stereoscopic camera disparity maps and was conveyed through the length of the sound files being played between 90-300ms duration - closer distances had briefer sonifications, referred to as its rhythm frequency (Bologna, Deville & Pun, 2010). Experiments have found the device useful for colour categorization (Bologna, Deville, Vinckenbosch & Pun, 2008) as well as navigation and object localization (Bologna et al., 2010).

2.4.1.5. Colour Enhanced Visual Sonifier, ColEnViSon (2009)

The ColEnViSon integrates visual saliency analysis with colour sonification (Ancuti, Ancuti & Bekaert, 2009). RGB space is converted to CIE LCH colour space (a cylindrical form of CIELab space) which contains 267 colour centroids based on the 'Universal Color Language and Dictionary of Names' (Kelly & Judd, 1976). The image is first colour-simplified by converting CIE LCH values to their nearest centroids (then adding luminance information). It then extracts and sonifies only isolated 'salient' regions in an image to reduce noise (selected using contrast irregularity on R-G and B-Y colour opponencies, orientation and mean-shift boundary-clustering techniques). The image is divided into larger pixels and 2x2 sections and a single column is sonified at a time. The Y-Axis is reflected in one of four musical layers / channels played simultaneously. The X-axis is scanned across over time. Ten colour categories are formed from hue and chroma, with additional luminance information given through higher pitched notes (across ten octaves) on the corresponding colour's timbre. Only the attack of the colour's timbre is played (the initial change from silence to peak volume levels, from which users can recognise the timbre and note), to reduce time and result in faster refresh rates and higher resolutions. With four musical layers and a 1/400 Tempo duration, one 64x64 image can be scanned through in 4 seconds.

Initial experiments compared the ColEnViSon to the vOICe in four novice users (30 minutes prior usage) and showed that the systems were comparable at detecting orientation and presence of complex stimuli. In this small sample, the ColEnViSon outperformed the vOICe on colour judgements (note that the vOICe only translates luminance), which demonstrates that colour information can be integrated without deterioration in spatial resolution.

2.4.1.6. EyeMusic (2012)

The EyeMusic sonifies a 24x40 pixel coloured image, typically over 2 seconds (Levy-Tzedek et al., 2012a, 2012b). Colours are first segregated into one of five categories (red, green, blue, yellow & white) and each category is encoded through timbre (e.g. white = piano, blue = marimba). The colour's vertical location is denoted through the pitch / note of the instrument, using notes across eight octaves of a pentatonic scale (higher pitches reflect higher spatial positions) while the brightness of each colour affects the loudness of the note (bright is loud and dark is quiet). Finally, each column is sequentially presented over time to complete the X-axis.

Initial experiments with the EyeMusic have focused on sensorimotor interactions between behaviours informed through either vision or the EyeMusic. Levy-Tzedek et al. (2012a) report a comparison between visually-informed reaching behaviour and SSD-informed reaching behaviour for a single white target object. Feedback for incorrect final hand positions was either on screen for the visual condition or sonified through a blue object for the EyeMusic condition. They found no significant difference between vision and the EyeMusic trials in users' movement time, peak speed and path length, indicating that EyeMusic's encoding could inform users' motor movements. However, there were significantly higher errors on endpoint localisation than with vision. A subsequent project by Levy-Tzedek et al. (2012b) examined the shared spatial representation of vision and EyeMusic soundscapes. Participants were required to indicate the location of an abstract target (either visually shown or heard using the EyeMusic) using a joystick. They found that altering the visual sensorimotor feedback (in this case, skewing the joystick's representation on screen by 30°) influenced not only the visual trial movements but also the EyeMusic trials (where no feedback was given). The former experiment indicates that soundscape feedback utilising timing, timbre, loudness and pitch can inform a spatial representation for movement, while the latter experiment indicates a shared spatial representation between vision and the EyeMusic soundscapes, as influenced by external feedback. While both of these experiments use simple target shapes on black (silent) backgrounds, limited colours and spatial target locations, these experiments demonstrate sensorimotor guidance that may be applicable to more natural complex environments where colour is meaningful rather than arbitrary.

Table 2.2. Coding strategies in current sound-colour SSDs.

SSD	Spatial Representation	Colour Representation	Audio-Colour Coding	Reference
SoundView	A single colour is selected via a stylus applied to a 2D tablet.	Hue, saturation and brightness.	White noise first goes through a low-pass filtering cutoff frequency denoting brightness[1]. Hue values determine which of 12 pitch-filters is applied[5] and this effect is modulated by saturation[1,4].	Doel, 2003.
Eyeborg	A single colour is made from a centrally-weighted distribution of pixels.	Hue and saturation.	Hue is either categorised into 12 or 360 microtones (1st version) or mapped to a continuous frequency between 363.797Hz and 717.591Hz (2nd version)[5]. Saturation is denoted through volume[4].	Else, 2012.
Kromophone	Single point. 2nd version codes the horizontal axis over time.	1st version uses hue, saturation and luminance. 2nd version uses red, green, blue, yellow and luminance.	1st version maps hue to pitch[5], saturation to left-right panning and luminance to volume[2]. 2nd version maps the contribution of red, green, blue, yellow and luminance to the volume of five unique sounds (each with unique panning, pitch and timbre[6]).	Capalbo & Glenney, 2009.
See ColOr	Horizontal axis coded via interaural timing differences. Depth coded through tempo. Global modules can select features for sonification.	Hue is transformed into 7 focal colours. In addition saturation and luminance are used.	The seven focal colours are given a unique timbre[6] each with their volume increasing nearer the colour's focal point. Increased saturation is communicated through higher pitched notes[3]. Finally darker colours add a double bass while brighter colours add a singing voice, with brighter and darker variants using higher and lower pitched notes respectively[1].	Bologna et al., 2007.
ColEnViSon	Horizontal axis over time. Vertical axis over audio channels. Bottom-up orientation modules select salient regions.	CIE LCH space is used (luminance, chroma and hue) to derive 10 focal colours.	Each colour is categorised according to its proximity to 'colour centroids' of focal colours. From this, unique timbres are given to 10 colours including greyscale[6].	Ancuti et al., 2009
EyeMusic	Horizontal axis over time. Vertical axis coded via note on 8 octave pentatonic scale (pitch).	Red, green, blue, yellow and white categories as well as a brightness dimension.	Each colour category is given a unique timbre[6] with the brightness of the colour denoted through its volume[2]. A colour's vertical position denotes its pitch on 8 octaves of a pentatonic scale[7]. A single vertical column is sonified with subsequent columns presented in a serial fashion over time.	Levy-Tzedek et al., 2012a, 2012b.

Supported as a correspondence by: [1]Ward et al., 2006; [2]Marks, 1987; [3]Orlandatou, 2012; [4]Giannakis, 2001; [5]Simpson et al., 1956; [6]Rossi et al., 2009; [7]Walker et al., 2009.

2.4.2. Potential future directions

Whereas many of the future directions for colour-to-touch depend on technological developments, the same does not apply to colour-to-sound as complex sounds are relatively simple to produce. Instead, it is suggested that future developments in colour-to-sound translation will primarily be based on finding optimal solutions with respect to human perceptual abilities.

The idea that there could be a law of association between colours and sounds (physical or psychological) has a very long history (Jewanski, 2010). In multi-sensory perceptual interactions, there are automatic associations between higher vertical locations with higher pitches (Melara & O'Brien, 1987; Roffler & Butler, 1967; Walker et al., 2010), as well as brighter luminances with higher pitch (Hubbard, 1996, Marks, 1974) and increased loudness (Lewkowicz & Turkewitz, 1980, Marks, 1987). These associations are exploited in many of the SSD devices described here. When asked to match colours with sounds, there is also a tendency to link saturation with pitch albeit non-linearly ('middle C' pitch is linked to the highest saturation, higher or lower pitches reduce saturation) and to timbre (pure tones are less saturated than instruments) (Ward, Huckstep & Tsakanikos, 2006). As for hue, one possibility is that this could be linked to various aspects of timbre. Timbre is formally defined as that quality of sound that enables different instruments to be distinguished (irrespective of loudness or pitch).

A higher-level approach to timbre is to get participants (music students, in this example) to rate sounds along different descriptive labels (Zacharakis, Pastiadis, Papadelis & Reiss, 2011). Using factor analysis they found three factors explaining 82% of the variance in timbre-descriptive terms: namely, volume / wealth (descriptions of 'full' and 'rich'), brightness / density ('brilliant,' 'deep,' 'bright' and 'thick') and texture / temperature ('soft,' 'harsh,' 'rounded,' 'warm' and 'sharp'). These orthogonal dimensions could potentially be mapped to red-green and blue-yellow dimensions to create a circular representation of hue.

In utilising perceptual aspects of timbre for the representation of colour, the most salient dimensions for human perception should be considered. In practice, timbre depends on the distribution of frequencies and how these change over time to create a spectral envelope. The spectral envelope for any given instrument timbre contains the 'attack' (change from silence to peak volume),' 'decay' (reduction of initial volume after peak), 'sustainment' (steady sound levels from holding a note) and 'release' (final reduction of sound levels to silence). Systematically varying these components offer new possibilities for the representation of colour.

Considering the distribution of frequencies in timbre, Von Bismarck (1974) identified four dimensions that explained 90% of the perceptual variance in artificially created steady-state timbres.

The two most influential dimensions were 'sharpness-dullness' (centre of gravity of the spectral distribution, dominant upper or lower harmonics), and 'compactness-scattered' (compact refers to harmonic combinations of frequencies while scattered refers to a wide selection of frequencies to create white noise). Since either end of these dimensions provide perceptual opposites (e.g. dominant upper harmonics is opposite to dominant lower harmonics) the use of these opposites can be used to represent opponent colours present in veridical colour perception such as yellow and blue. In addition, if the other dimension is positioned as a perpendicular axis, this provides another ability to represent opponent colours (e.g. red and green) through perceptually opposite sounds (harmonics and noise). In combination this soundspace provides neighbouring sounds that are more similar than their opponencies, which mimics the structure and position of opponent colours in veridical colour perception.

Considering the temporal dynamics of timbre (attack-decay-sustainment-release profile), Wessel (1979) found, using multi-dimensional scaling, that the temporal dimension was mostly defined by instrument family, noting three distinct groups (Trumpet-trombone-French horn; oboe-bassoon-clarinet; violin-viola-cello). Grey (1975, 1977) found that three dimensions could explain the majority of variation in dynamic timbre (with equalized volume, pitch and duration), these being the overall spectral energy distribution and, over time, the synchronicity of upper harmonics (where high frequency sounds all increase and decrease in volume together) and low energy high frequency noise during the attack segment (subtle disharmonic content present just as the note is played). Barrass (1997) used Grey's timbre dimensions to create 3D space containing instrument sounds at specific locations so that similar sounds would be spatially close together and dissimilar sounds would be far apart. The purpose of this was to mimic the structure of HSL colour space but using sounds instead of colours. He matched luminance-pitch (linear), saturation-'upper/lower overtone emphasis' (linear) and hue-timbre (circular) dimensions to create a perceptually distinctive information sound space. In order to construct a circular representation of hue from timbre, two opponent axes (similar to RG and BY axes in CIELab space) were created from 'upper harmonic synchronicity' and 'high frequency noise during the attack phase' to select the appropriate instruments (Grey, 1975). As a result, the relationship between two colours (e.g. are they perceptually close or far apart?) could be understood through how the sounds from those colour locations relate to one another in terms of perceptual similarity or dissimilarity.

The ability to discriminate two sounds from one another is an essential characteristic in determining that these sounds could represent different properties, such as two colours that differ from one another. But while the sounds might indicate that 'sound A' represents something different than 'sound B,' it does not by itself indicate what specific shades of colour sounds A and B might represent. One way that you can imply that certain sounds represent certain colours is through using pre-existing intuitive associations between sounds and colours in your cross-modal mapping. One way of establishing any cross sensory is simply through asking large groups of participants which colours they think would 'best suit' a variety of timbres. Rossi, Perales, Varona and Roca (2009) explored the colour associations found for instruments (played in a melody). They note that some associations occur more often than expected by chance (e.g. guitar=red, vibraphone=yellow). Currently, there is no indication as to why these associations would be this way around, so further investigations should seek to understand the rationale behind these choices made by participants and whether or not the use of these affects individuals' ability to learn colours through sound.

2.5. Summary and conclusions

This paper provides the first overview of colour representation through tactile and auditory modalities, which reveal a wide variety of methods for communicating different types of colour space. It has been shown that colour information aides in figure-ground segmentation, scene recognition and object identification, especially within low resolution vision (Torralba, 2009). As a result, traditional greyscale SSDs may benefit more from additional colour dimensions rather than just increases in resolution. By examining both human factors and solutions used by SSDs in representing colour, this can inform us on appropriate modality for a given task. Finally, as tactile and auditory feedback are not mutually exclusive outputs, they allow us to justify approaches to hybrid tactile-audio visual devices.

In choosing the appropriate modality for colour representation, the first consideration is the level of information to be encoded. Colour can be simple (e.g. symbolic coding, EyeMusic), perceptually accurate (e.g. CIE LCH, ColEnViSon) or somewhere between the two (e.g. RGB, VIDET glove; HSB, Soundview), this information content being magnified by both the spatial axes (i.e. single point, horizontal, vertical, depth) and spatial resolution. Finally, when it comes to communicating this information, is this done as a whole (e.g. See ColOr), piecemeal (e.g. EyeMusic), manually selected (e.g. Soundview) or computationally selected (e.g. ColEnViSon). For computationally selective devices, users need feedback on why / where objects are selected, otherwise the users' 'visual world' risks becoming unpredictable, shifting the users' focus to the tool rather than the tool's use (Dotov, Nie & Chemero, 2010). This provides us with the amount of information that needs to be transferred at a given moment. Kokjer (1987) notes that when information is presented sequentially, auditory input has a higher and faster discrimination rate than comparable tactile stimulation. In contrast, the skin provides more stimulation points at a given time. This implies that high resolution information is best presented serially to the auditory channel while low resolution information is presented in parallel to the tactile channel, with a combination being an ideal choice for a hybrid device. We can see these effects from higher spatial resolutions possible through auditory SSDs (Striem-Amit et al., 2012) while higher temporal resolutions (required for motion perception) are possible with tactile SSDs (Matteau, Kupers, Ricciardi, Pietrini & Ptito, 2010). A combination of both mirrors our own visual system's parvocellular and magnocellular routes best suited to detail and motion respectively.

If the device is suitable for wide-spread long-term use, blind user preferences should be considered (Golledge, Marston, Loomis & Klatzky, 2004). A minimal or intuitive use of the sense organs that blind individuals rely on day-to-day is preferable. While continuous use of the fingers, ears and tongue is understandable for short-term experimental use, this is likely to significantly hamper environmental manipulation, important auditory-cues and social interaction for blind individuals. Colour in of itself is unlikely to be a complete SSD solution for many blind individuals, so complementary information for navigation or human recognition may be desirable to facilitate long-term use (Khan, Moideen, Lopez, Khoo & Zhu, 2012; Mann et al., 2011). Expanding SSD use may also rely on utilising commercially available devices with video-input and audio / tactile output (such as tablets / smartphones) and allowing easy access to SSD programs. Unfortunately, many colour SSDs lack these features, confining their use and potential scientific impact, irrespective of their utility. The wide variety of directions taken in the design of SSDs can be seen to reflect another problem of failing to compare devices across common benchmarks. This makes establishing which principles of design work and why difficult to ascertain. Uses of colour that are natural and universal (e.g. fruit identification) may prove to be ecologically valid and fair tests of such devices.

Taking into account the selected modalities' perceptual limitations and cross-modal correspondences should help inform the tactile / auditory SSD output. Firstly, a consideration of perceptual resolution is required, considering the difficulty in separating frequency and amplitude (Harris, Arabzadeh, Fairhall, Benitro & Diamond, 2006; Taylor, 1977; Verrillo et al., 1969), limitations in spatial discrimination (Grantham, Hornsby & Erpenbeck, 2003; Kaczmarek, Webster, Bach-y-Rita & Tompkins, 1991), adaptation rates (Hollins et al., 1990) and how these vary across individuals (Morley & Rowe, 1990; Roy & Hollins, 1998). A consideration of these factors allow for optimal psychophysical tuning of the output to reflect the user's ability to discriminate information. As a last touch, pre-existing cross-modal correspondences could assist in linking dimensions of colour to sound / touch in an intuitive manner. The final structure of 'colour' space created by this process can be compared to veridical colour perception through similarity judgements between colours (e.g. Kahol et al., 2006).

Colour SSDs that use alternative approaches to colour space (i.e. linear, symbolic or single point) have been discussed alongside their limitations in replicating colour space. For instance, linear representations of hue produce unintuitive mathematical relationships (e.g. blue may be double the red frequency), lack both a smooth violet-red transition and focal points. Symbolic representations lack the transitional aspects of colour and translating single points of colour means that establishing figure-ground segmentation and scene / object recognition must be done piecemeal (Torralba, 2009).

Finally, colour SSDs also provide the ability to investigate modality-independent theories of visual perception (O'Regan & Noë, 2001) or shared processes (Levy-Tzedek et al., 2012b). As SSDs replicate more and more of the features of veridical colour receptors, human-environment colour interactions and the structure of human colour space, comparable colour knowledge and even perception should be possible. Each of these rules for colour-interactions is alterable and provides the ability to dissect the necessary requirements for 'colour-like' experiences. Finally, the ability to examine the long term effects of colour SSDs are rare (e.g. NH's EyeBorg experience) but are required for future investigations.

3. Tactile-Vision Synaesthesia: Tactile Discrimination and Phenomenology

3.1. Abstract

A questionnaire was administered (N = 21) to self-reported tactile-vision synaesthetes to find commonalities in their reported phenomenology. It was found that emotional human touch, itchiness and sexual experiences were the most common inducers; suggesting that affective processing may underpin this in many cases. These inducers are discussed with reference to common underlying neural mechanisms such as the insula. Three tactile-vision synaesthetes underwent tactile-colour consistency testing to inanimate objects felt by the hand. All showed higher consistency as a group, but only one individual reached significance relative to controls. During consistency testing, controls showed a weight-luminance association (heavier objects are judged darker), and interestingly, this is also present for synaesthetic photisms. Three psychophysical discrimination tasks were also administered, examining their spatial discrimination of tactile texture, tactile orientation discrimination and tactile-visual integration. The synaesthetic group had a significantly enhanced spatial discrimination of texture relative to controls. However, their orientation discrimination and tactile-visual integration was comparable with controls. It is suggested that the enhanced spatial discrimination of texture for synaesthetes is due to the unique recruitment of visual processing regions not found in sighted controls. Since the orientation and tactile-visual integration tasks already use visual and shared processing regions for controls the synaesthetes have no unique advantage here. Finally, a task exploring the psychophysics of photisms in relation to touch is reported, finding a close spatial mapping between the location of touch and colour for synaesthetes. This is discussed in relation to shared spatial representations for touch and vision. These findings enhance our understanding of the tactile-visual system from the largest group of tactile-vision synaesthetes gathered to date.

3.2. Introduction

Processing regions normally associated with vision have been demonstrated to show surprising flexibility to interact with information obtained through touch. These include the functional recruitment of ‘visual’ regions for tactile tasks in sighted individuals (Zhang et al., 2005), intuitive tactile-visual mappings (Ludwig & Simner, 2013; Walker, Francis & Walker, 2010), the unmasking of tactile routes deep into the visual cortex after visual deprivation (Merabet et al., 2008) and even somatosensory sensations from occipital cortex stimulation (Kupers et al., 2006; Ptito et al., 2008). These explorations into tactile-visual interactions may provide a basis for predicting when and how the production of visual phenomenology from touch can occur in developmental or acquired tactile-vision synaesthesia (Armel & Ramachandran, 1999; Simner & Ludwig, 2012). In this introduction, we will first summarise the current evidence surrounding tactile-vision synaesthesia and its link with intuitive mappings between touch and vision found in the general population. Then we will look at how 'visual' brain regions serve a functional role in seemingly tactile-only tasks. Finally we will examine the neural changes associated with synaesthesia and what this implies for their performance on a variety of tactile discrimination tasks. This provides the framework for the experimental studies that follow.

3.2.1. Tactile-vision synaesthesia: Previous investigations and consistency testing

Synaesthesia is a cognitive condition that affects approximately 4.4% of the population (Simner et al., 2006) in which stimulation in an inducing modality (e.g. tactile stimulation) creates an automatic, consistent and explicit experience of a different concurrent modality (e.g. vision) that is not found in the wider population (Ward, 2013). Tactile-vision synaesthesia is a relatively rare form of synaesthesia present in 6 to 9% of synaesthetes, and co-occurs to a larger degree with ‘coloured sensation’ synaesthesias, such as pain-vision or temperature-vision synaesthesia (Novich, Cheng & Eagleman, 2011). Initial investigations into tactile-vision synaesthesia have found that varying tactile stimulation can systematically influence the subsequent visual experience (Smith, 1905), noting "Hard objects... are dark in color; soft objects are light" (p. 260). More recent studies have examined how properties of touch (e.g. smoothness) and vision (e.g. luminance) are aligned in both synaesthesia and the wider population (Ludwig & Simner, 2013; Simner & Ludwig, 2012; Ward, Banissy & Jonas, 2008). However, the number of such cases is small and one tactile-vision synaesthete, EB, has undergone a consistency test in order to support claims of genuineness (Simner & Ludwig, 2012). Simner and Ludwig (2012) report that touch-colour pairings to simple haptic stimuli (e.g. objects varying on one dimension like roundness) did not mark EB as more consistent, but complex stimuli (e.g. objects varying unpredictably in shape, softness and texture) did. The reasons behind this discrepancy include high levels of touch-colour consistency in controls for simple tactile stimuli, where utilising both intuitive tactile-vision mappings (e.g. Smoothness-luminance) and implied real world colours (e.g. Wood-brown) allow non-synaesthetes to use these rule sets to aid consistency. By contrast, the more difficult to recognise complex tactile stimuli can make the non-synaesthetes' use of these rule sets harder to implement, leading to more colour-inconsistency. Another factor to consider is that these tests are designed according to the inducers of a single tactile-vision synaesthete, so it is not clear if these methods are suitable for all tactile-vision synaesthetes or even whether consistency is a widely reported aspect of this synaesthesia.

3.2.2. Tactile-visual interactions in discrimination

Besides the intuitive mappings between touch and vision seen above, tactile-visual interactions are also crucial for certain forms of tactile discrimination processing. Some tactile discrimination tasks such as with *passively* obtained texture (a single occurance of spatially distributed variations of pressure, as opposed to *actively* explored surfaces by an individual's hand which involves motion or stroking) retain functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET) activation patterns within the primary (S1) and secondary (S2) somatosensory cortices as well as the right angular gyrus (RAG) – henceforth S1, S2 & RAG (Sathian, Zangaladze, Hoffman & Grafton, 1997; Zhang et al., 2005). Further to this, raising the neural firing threshold in the RAG through offline anodal tDCS (transcranial direct current stimulation administered before the task) can improve tactile distance judgements similar to passive texture discrimination (Spitoni et al., 2013). Discriminating tactile textures can also functionally recruit visual regions such as V1 during visual deprivation, an area well suited to spatial-distance evaluations (Merabet et al., 2004, 2007, 2008). Interestingly, variations in task demands such as with an orientation test of tactile acuity or tactile motion processing activate brain regions typically associated with orientation and motion processing in vision (Blake, Sobel & James, 2004; Galletti, Battaglini & Fattori, 1991; Hagen, et al., 2002; Sathian et al., 1997; Zhang et al., 2005). For tactile orientation discrimination, 'visual' regions such as the parieto-occipital cortex (POC) display a functional role in their processing so that when they are inhibited by TMS stimulation, performance on orientation but not passive texture discrimination revert to chance performance (Zangaladze, Epstein, Grafton & Sathian, 1999; Zhang et al., 2005). In contrast, inhibiting earlier processing stages such as the S1 impair both orientation and texture discrimination (Zangaladze et al., 1999). Further supporting S1's role in the gratings orientation task, enhanced tactile acuity has been observed by increased attention and involvement of frontal mechanisms (Dionne, Meehan, Legon & Staines, 2009) and through lowering S1's activation threshold through offline anodal tDCS (Ragert, Vandermeeren, Camus & Cohen, 2008). All this points to several key facts about texture and orientation discrimination tasks: firstly that the S1 is required for both; secondly that enhancing excitability in S1 can improve orientation performance (changes to texture performance are unknown); and finally that the POC is functionally recruited for orientation discrimination while the RAG is functionally recruited in passive texture discrimination. For touch-colour synaesthesia, one intriguing possibility is that the process for discriminating distances or orientation information in the RAG and POC respectively might further influence the visual photisms elicited. If so, this could be indicative of the routes used in eliciting colour from touch in the brain.

Most evidence on tactile-visual texture integration involve the *active* exploration of surface properties, this active exploration is quite different from passive stimulation where the tactile stimulus is pressed into the skin, for a variety of reasons. While passive texture involves distance judgments between points of pressure, active exploration of texture involves additional vibrational, muscle resistance and softness cues gathered from exploratory movements to evaluate surface 'roughness' (Lederman & Klatzky, 2009). These active explorations of texture roughness involve a wider network of regions outside of somatosensory regions than those seen during passive texture discrimination (Simões-Franklin, Whitaker & Newell, 2011). Stilla and Sathian (2008) report an fMRI study into shared regions during active exploration of tactile texture and the identification of its visual equivalent, finding a shared activation in the medial occipital cortex (MOC). A subsequent multivariate Granger causality analysis on this data suggested a bottom-up route whereby tactile processing meets visual texture-sensitive processing after S1 and insula involvement (Deshpande, Hu, Stilla & Sathian, 2008). Despite the evidence for shared neural processes, the behavioural evidence supports the view that tactile and visual texture perception are treated as independent sources of information rather than being integrated across the senses (Guest & Spence, 2003; Whitaker, Simões-Franklin & Newell, 2008). For example, roughness judgments are primarily determined through touch, and visual roughness information does not bias or improve this discrimination, instead reduced discrimination is seen as a result of dividing attention (Guest & Spence, 2003b). Likewise in bimodal tasks, the preference given to tactile or visual stimulation typically reflect task demands, with evaluations of 'spatial density' (i.e. the distance between raised dots) favouring visual information and 'roughness' (i.e. the perceived resistance to active exploration) favouring tactile information (Lederman, Thorne & Jones, 1986). While this lack of integration is surprising for controls, it provides a potential route for tactile-visual mappings in synaesthesia. For instance, if the MOC provides a shared area for tactile and visual texture processing, then will tactile-vision synaesthetes' photisms display a visual texture comparable with that of the inducing tactile texture? Furthermore, will the structural and functional changes associated with synaesthesia influence the integration between tactile and visual texture information (Rouw, Scholte & Colizoli, 2011)?

3.2.3. Neurological and perceptual changes in synaesthesia

Previous research on more common forms of synaesthesia such as grapheme-colour have indicated increased levels of neural activation, grey matter density and white matter tract coherence in locations related to the individual synaesthetic experience (Banissy et al., 2012; Barnett et al, 2008b; Rouw, Scholte & Colizoli, 2011; Terhune, Tai, Cowey, Popescu & Cohen Kadosh, 2011). In addition, the modalities recruited by synaesthesia appear to be bound together by parietal regions (Esterman, Verstynen, Ivry & Robertson, 2006; Muggleton, Tsakanikos, Walsh & Ward, 2007). Beyond neurological changes, synaesthetes have also been observed to have increased tactile and colour acuity in synaesthetically affected *concurrent* modalities (Banissy, Walsh & Ward, 2009). For instance, in mirror-touch synaesthesia, where viewing other individuals being touched results in illusory touch on the synaesthete's own body, Banissy et al. (2009) found improved tactile discrimination using the JVP gratings orientation test, which requires contributions from both the S1 and POC (Zangaladze et al., 1999; Zhang et al., 2005). Increased activation in early S1 regions has been found to correlate to the misattribution of touch on the self in non-synaesthetes and mirror-touch synaesthetes (Blakemore, Bristow, Bird, Frith & Ward, 2005; Keysers, Kaas & Gazzola, 2010). Similarly, artificially increasing excitability in the primary somatosensory cortex can also lead to similar increases in performance on the JVP gratings orientation test (Ragert et al., 2008). Therefore it seems likely that hyperactivity in S1 can account for the discrimination improvements seen in mirror-touch synaesthesia alone, rather than improved S1 to POC information transfer or improved POC discrimination abilities. In contrast, for *inducing* modalities, enhanced performance has been observed in temporal discrimination tasks only when concurrent modalities are able to assist in their processing (Saenz & Koch, 2008). As no enhanced discrimination has been reported for modalities that only *induce* other percepts, it would not be expected for tactile-vision synaesthetes to have increased sensitivity within S1 unlike with mirror-touch synaesthetes for whom it is their concurrent. However there may be increased involvement of the RAG or POC which would affect tactile texture and orientation tasks, particularly for synaesthetes who show distance or orientation information in their photisms. One possibility is that tactile-vision synaesthetes may have additional information transfer between S1 and the RAG / POC or more hyperactive regions associated with these tasks, which would lead to better tactile discrimination abilities.

There are a variety of tactile-visual routes demonstrated above that may serve as a basis for tactile-visual synaesthesia. By exploring phenomenological and behavioural evidence that draw upon these routes, it becomes possible to gauge which of these processes are used or altered through the presence of synaesthesia. In our behavioural experiments, we expect higher touch-colour consistency than controls, as well as improved performances on the tactile texture, tactile orientation and tactile-visual integration tasks due to the additional visual processing important to the completion of these tasks.

3.3. Experiment 1 - Questionnaire

Many aspects of tactile-vision synaesthesia remain unclear, from the phenomenology of those with the condition, to their touch-colour mappings and any effects this tactile-visual coupling might have on sensory discrimination. Much of this can be ascribed to its rarity, with Day (2014) reporting that self-reported touch-colour synaesthesia makes up 4.07% of a sample of 1007 synaesthetes. Taken alongside the 4.4% prevalence of synaesthesia in the wider population (Simner et al., 2006), we would expect to see an incidence of one case per 558 people. Due to the rarity of this form of synaesthesia we used an online questionnaire in order to reduce the barrier for entry so that a more representative sample can be obtained. From this questionnaire it is hoped that the phenomenology of this synaesthesia can be better understood and that subsequent behavioural testing can utilise this information to examine variations or subtypes of these synaesthesias.

3.3.1. Participants

All respondents to the questionnaires were recruited through advertisements on synaesthesia e-mailing lists, focused on both the UK (UK Synaesthesia Association) and internationally (www.daysyn.com). No formal test to prove the presence of synaesthesia is required to be a member of either list. For the tactile-vision synaesthesia questionnaire, twenty-one respondents reported experiencing synaesthetic visions from tactile stimulation (5 male, 16 female), with a mean age of 32.43 (SD = 12.07). Participants were informed that all responses were both anonymous and confidential.

3.3.2. Materials and procedure

Due to the rarity of previously published work on the topic of tactile-vision synaesthesia, questions were based on previous case studies on tactile-vision synaesthesia (Simner & Ludwig, 2012; Smith, 1905; Ward et al., 2008), some questions were more exploratory relating to variations in sensory stimulation and finally revisions were made based on feedback from a pilot questionnaire given to a single tactile-vision synaesthete. The questions are listed in Table 3.1 with the question number indicating the original order in the questionnaire. All questions are rated by respondents on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5), with the exception of location based questions (i.e. question 4) which had a 5-point scale ranging from never (1) to always (5) experiencing synaesthesia from that area being stimulated. All questions are accompanied by open text boxes that participants could comment on or clarify their responses. The questionnaires were administered online using the Bristol Online Surveys service.

3.3.3. Results and discussion

Descriptive statistics for each of the questionnaires are provided in table 3.1. All statistics are ordered by section, with the ordering within each section done via average agreement by the synaesthetes with the highest agreement at the top, lowest at the bottom. Questions with very high levels of agreement (>4) or disagreement (<2) are highlighted in bold for ease of reference. These descriptive statistics are followed by quotations given by synaesthetes during the completion of these questionnaires. While it could be argued that describing the questionnaire results by way of means and standard deviations might be misleading sine it presumes the categories given for participants to choose to be similar to that of a continuous scale (Kostoulas, 2013), even if this were the case, this information is informative to the distribution of responses, and from this general tendencies (Lord, 1953). In addition, other work has found this treatment to yield no substantial or statistical difference over ranked approaches (de Winter & Dodou, 2010).

Table 3.1. Tactile-vision synaesthesia questionnaire responses

Qu. 1. When do you experience this synaesthesia?	Mean	SD
Qu. 1.5 - When I am touched by an emotionally meaningful person to me	**4.38**	0.97
Qu. 1.4 - When I am touched by a stranger	3.86	1.06
Qu. 1.2 - If an inanimate object bumps into me	3.29	1.55
Qu. 1.3 - If I touch my own skin	3.20	1.47
Qu. 1.1 - If I touch an inanimate object	3.19	1.50
Qu. 1.8 - All touch sensations	2.85	1.46
Qu. 1.6 - Viewing others touching each other (such as a handshake or hug)	2.43	1.50
Qu. 1.7 - Seeing others touch an object I am using	2.40	1.47
Qu. 2. Which aspects of touch can create visual experiences?		
Qu. 2.9 - Feeling itchy	**4.21**	1.27
Qu. 2.10 - Sexual experiences	**4.19**	0.93
Qu. 2.7 - Objects pressing into me	3.65	1.57

Qu. 2.8 - Objects brushing past me lightly	3.63	1.50
Qu. 2.1 - Holding a vibrating object	3.50	1.47
Qu. 2.2 - Holding a pointy object	3.25	1.41
Qu. 2.3 - Holding a heavy object	3.20	1.47
Qu. 2.4 - Holding a rounded object	3.05	1.54
Qu. 2.6 - Holding a light object	3.00	1.49
Qu. 2.5 - Holding an object that doesn't move	2.80	1.51
Qu. 3. Are these aspects of touch reflected in your vision?		
Qu. 3.4 - The intensity of touch is the intensity of my visual experience	3.26	1.45
Qu. 3.5 - The location of the touch on me reflects the location in space I see the vision	2.67	1.46
Qu. 3.2 - The texture I feel is the texture of my visual experience	2.57	1.57
Qu. 3.3 - The softness I feel is the softness of my visual experience	2.57	1.40
Qu. 3.1 - The shape of the tactile sensation is the shape of my visual experience	2.55	1.67
Qu. 3.6 - When I see others touching, my visual experience appears near them	**1.86**	1.15
Qu. 4. Where being touched on your body can produce visual experiences?		
Qu. 4.1 - Fingers	3.71	1.10
Qu. 4.7 - Face	3.60	1.23
Qu. 4.2 - Palms	3.50	1.19
Qu. 4.4 - Shoulder	3.50	1.28
Qu. 4.6 - Back	3.40	1.39
Qu. 4.3 - Arms	3.35	1.35
Qu. 4.10 - Legs	3.32	1.42
Qu. 4.8 - Tongue	3.24	1.58
Qu. 4.11 - Feet	3.24	1.51
Qu. 4.12 - Toes	3.10	1.55
Qu. 4.9 - Hips	3.05	1.50
Qu. 4.5 - Chest	3.00	1.52
Qu. 5. Does varying these experiences also vary the visual experience?		
Qu. 5.7 - Different textures (i.e. smooth or rough)	**4.10**	1.07
Qu. 5.5 - My mental state (if I am tired or alert)	3.86	1.24
Qu. 5.4 - The length of time I am touched (i.e. brief or prolonged touch)	3.67	1.24
Qu. 5.2 - The type of touch I experience (i.e. light brush or tight grasp)	3.62	1.16
Qu. 5.9 - If I close my eyes (vs seeing the touch)	3.52	1.60
Qu. 5.6 - If I have alcohol vs being sober (don't answer if you don't drink)	3.50	1.38
Qu. 5.3 - The enjoyment (or dislike) of the touch I experience	3.48	1.36
Qu. 5.1 - Where I am touched (i.e. arm or leg)	3.19	1.03
Qu. 5.8 - Being touched in an area I cannot see (i.e. on my back or on my front)	3.10	1.58
Qu. 6. What types of visual content do you experience from touch?		
Qu. 6.2 - The visual content has colour	**4.33**	1.24
Qu. 6.4 - The visual content has motion (the content shifts around in space)	**4.00**	1.45
Qu. 6.5 - The visual content has a texture (if the visual content contains patterns / features)	**4.00**	1.05
Qu. 6.7 - The visual content appears within my mind's eye	3.90	1.37
Qu. 6.3 - The visual content contains simple shapes	3.52	1.44
Qu. 6.8 - The visual content appears at the location I am touched	3.14	1.28
Qu. 6.1 - The visual content has black and white	2.76	1.81
Qu. 6.6 - The visual content appears out 'in the world'	2.40	1.60

Touch by emotionally significant people appears to be the strongest inducer of synaesthetic visions (4.38), since this is lower for strangers (3.86), self-touch (3.20) and touching objects (3.29, 3.19). This indicates that the touch sensation itself may not be the whole picture, instead emotional affect may play a stronger role. This could be the effect of emotion modulating the somatosensory cortices (Damasio et al., 2000), it could be directly related to emotional processing itself, or the increased attention given to emotive tactile stimulation. One synaesthete noted that their synaesthesia "...tends to increase in intensity when I am tired, or my skin is particularly sensitive", while tiredness seems to affect photisms to a moderate degree (3.86), increased tactile sensitivity is explicitly mentioned by some respondents as modulating their photisms. In support of emotional processing and attentional focus is the surprising finding that the most powerful inducer appears to be that of 'itchiness' (4.21), a sensation that is known to have a strong affective and attentional component (Valet et al., 2008). Interestingly, all of the most powerful inducers such as itchiness (4.21), emotional touch (4.38) and sexual experiences (4.19) are involved with similar neural regions, specifically through increased insula activation. The insula is a region correlated with processing itch (Herde, Forster, Strupf & Handwerker, 2007; Holle, Warne, Seth, Critchley & Ward, 2012), emotional touch (Morrison, Björnsdotter & Olausson, 2011; Olausson et al., 2002), intensity of touch (Lovero et al., 2009) and orgasm (Komisaruk et al., 2004).

Inanimate objects, whether actively touched or passively received by the synaesthete appear to be a source of disagreement between synaesthetes as an inducer of photisms (3.29, 3.19). This indicates that previous experiments looking at consistency testing for tactile-vision synaesthetes may only be able to meaningfully ask the question to a particular subset of synaesthetes (Simner & Ludwig, 2012; Ward et al., 2008). Consistency testing is further compounded by some remarks made by the synaesthetes themselves, "I don't always experience it, nor is it always consistent. For example, brushing the back of my hand will not always produce the same colors or patterns." As such, consistency testing as a method of supporting the presence of synaesthesia may not be the optimal solution for confirming all variations of tactile-vision synaesthesia (Simner, 2012). The visual phenomenology is reported to contain colour, motion and texture (4.33, 4.00, 4.00). Visual texture can refer to the look of surface properties (e.g. 'wood grain') or more abstract patterns like Klüver's form constants (e.g. lattices). One synaesthete reported that many of her photisms from tactile stimulation involved these form constants, while another reported spatial stimulation directly affecting their visions "I see every groove, line, and bump when I touch the wallpaper." If colour is not always consistent for these synaesthetes, an exploration of motion or texture-consistency may be another factor to consider in confirming the presence of synaesthesia (Eagleman & Goodale, 2009).

Variations in touch do not appear to be uniformly affected by shared linguistic terms such as 'intensity' or 'softness' (3.26, 2.57). Photisms do not uniformly appear to occur at the location of the touch (2.67), and varying the area of tactile stimulation also tends not to universally affect the photisms (3.19). As might be expected from this, tactile stimulation to all locations on the body have a similar level of agreement for eliciting photisms (3.71 to 3.00), of note is that regions with a larger cortical representation on the homunculus (i.e. fingers, face) are reported to elicit photisms slightly more often.

The perceived texture or shape of tactile stimulation does not appear to be directly translated into their visual photisms (2.57, 2.55). However variations in texture do vary the content of the visual content in a more abstract way (4.10), a personal account of variations in synaesthetic photisms to different textures can be seen in figure 3.1. Exploring texture usually refers to information gained from active tactile exploration (Lederman & Klatzky, 2009). In agreement with previously described common inducers is that active exploration of different textures has been found to involve insula activation (Deshpande et al., 2008).

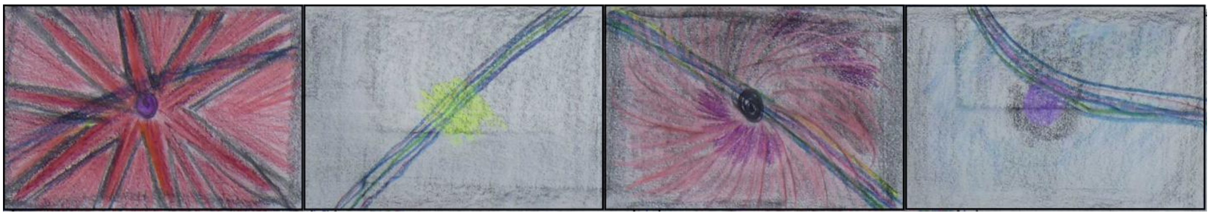


Figure 3.1. (Photism-illustration – roughness texture) Drawings of synaesthetic photisms kindly provided by PM who was unavailable for further behavioural testing. Left to right boxes display visual experiences to fake fur, micro suede, coarse sandpaper and smooth wood grain.

The location of the visual phenomena appear to include more 'associators' (3.90) than 'projectors' (2.40), illustrating this some synaesthetes report explicitly internal representations "I both see it in my head and feel the texture and the shape of the touch in my whole body" while others are more external "Moving my hand through the air and feeling the air rush past produces a see-through-like flame that sits on my hand as long as the air is felt." Other phenomenological reports of 'associator' and 'projector' tactile-vision synaesthetes were previously described by Ward, et al. (2008). This data falls in line with the previously observed spatial subtypes of synaesthesia (Rouw & Scholte, 2010; Ward, Li, Salih & Sagiv, 2006). Finally, although this was not addressed in the present research, some synaesthetes mentioned the length of time that these visions persist past the point of initial tactile stimulation, "visions of moving coloured shapes occur in sequences lasting from a few seconds up to 10 minutes. The sequences change often and may have only vague continuity from one to the next... there is often no apparent cause for new colours to come" (Steen, 2001). These phenomenological self-reports introduce further problems for consistency testing with photisms to the first stimuli persisting over subsequent stimuli.

Overall, touch-colour synaesthetes report varying levels of automaticity, consistency and what tactile stimuli induce colours. These variations in phenomenology may reflect different mechanisms or subtypes (e.g. consistent vs inconsistent pairings) and so a careful consideration of an individual's phenomenology may lead to different optimal techniques for 'confirming' the presence of synaesthesia. So far the only available data-point for consistency is from Simner and

Ludwig's (2012) case study, who found a tactile-vision synaesthete to be more consistent than controls on complex tactile stimuli. Whether this behavioural touch-colour consistency applies to a larger group of synaesthetes will be further explored in our behavioural testing.

3.4. Experiment 2 - Behavioural Testing

It is currently unknown whether touch as an inducing modality reflects a behavioural profile similar to mirror-touch synaesthesia, where touch is a concurrent (and associated with enhanced tactile sensitivity – Banissy et al., 2009), or whether touch-colour synaesthesia has its own pattern of behavioural characteristics of touch, more related to the additional visual processing that occurs. The discrimination abilities of touch-colour synaesthetes can help to identify which processing routes are altered through the presence of this particular variant of synaesthesia. This can help to identify not only the perceptual and behavioural changes associated with this type of synaesthesia but help inform future studies on its likely underpinning neurological profile. Touch-based tasks are known to utilise a wide variety of brain regions (including some of those typically linked to visual perception) depending on the task being performed. Changes to sensitivity within these regions can influence task performance (Ragert et al., 2008), an attribute previously found in synaesthesia (Banissy et al., 2009). By comparing synaesthetes and controls on a variety of tasks that are known to load upon different regions, which systems are affected and potentially used in synaesthesia can be identified. Texture tasks have been found to involve S1, S2, RAG and in the visually deprived, V1 regions (Merabet et al., 2004, 2007, 2008; Sathian et al., 1997; Spitoni et al., 2013; Zhang et al., 2005). Tactile orientation tasks involve S1, S2 and the POC while tactile-visual texture comparison tasks incorporate the S1, insula and MOC (Deshpande et al., 2008; Stilla & Sathian, 2008; Zangaladze et al., 1999; Zhang et al., 2005). Finally a photism-illustration task was given to explore how variations in spatial stimulation on the skin are related to synaesthete's photisms.

3.4.1. Participants

The tactile-vision synaesthesia group was comprised of three participants (MLS, CS & ES – See table 3.2.) with self-reports of consistent visual experiences from touching inanimate objects. Mean age was 41.67 (SD = 23.71), all participants were female, 1 participant was left handed and the highest level of education ranged from A-levels to postgraduate degree. Participants were recruited from the UK Synaesthesia Association's e-mailing list, Sean Day's International Synaesthesia List and the Sussex Synaesthesia database. All participants reported having sound-colour and pain-colour synaesthesia alongside a variety of other synaesthesias (see table 3.2), only MLS reported tactile concurrents with mirror-touch synaesthesia. The control group was comprised of 31 participants with no reported synaesthetic experiences, mean age was 35.06 (SD = 21.34), five participants were male, two participants were left handed and the highest level of education ranged from A-levels to postgraduate degree. One control participant was excluded for reported hallucinatory tactile sensations and relatives with grapheme-colour synaesthesia. A comparison of the ages of the two groups using a two-tailed Mann-Whitney U Test, which yielded a non-significant difference, $U(34) = 26.50$, $Z = -1.22$, $p = .221$. Travel expenses were reimbursed with students being given course credits for participation while non-students received £10 payment. Ethical approval was obtained from the local research ethics committee from the University of Sussex. Due to technical errors, four tactile-visual texture comparison scores had to be omitted for controls.

Table 3.2. Tactile-vision synaesthetes

Subject	Age	Sex	Handedness	Other synaesthesias	Touch as a concurrent
MLS	22	F	L	G-C, S-C, P-C, M-T	Yes
CS	68	F	R	G-C, S-C, P-C, Te-C	No
ES	35	F	R	S-C, P-C, Te-C	No

Key: M = Male, F = Female, L = Left, R = Right, G-C = Grapheme-colour, S-C = Sound-colour, P-C = Pain-colour, Te-C = Temperature-colour, M-T = Mirror-touch.

3.4.2. Design

A between-subjects design was used to contrast synaesthetes and controls across four tasks; touch-colour consistency, tactile-texture discrimination, tactile-orientation discrimination and tactile-visual texture discrimination. For the consistency task, the dependent variable was the amount of test-retest colour error, with each synaesthete being assessed individually against the control group. For the discrimination tasks, the independent variable was the group (synaesthete, control), while the dependent variables were either the percentage correct score for the tactile and tactile-visual texture tasks or the worked out hypothetical grating gap width (in mm) that would produce a 75% correct response rate for the tactile-orientation task (smaller mm gaps indicate better tactile discrimination). These were analysed using a series of Mann Whitney U tests due to the low sample size and uncertainty about normal distribution for the synaesthetic group.

3.4.3. Materials and procedure

A custom testing booth was created to allow access to the experimental laptop and provide comfortable visual-obscuring of both the participant's arms and the experimenter when they are in the testing position. Participants sat approximately 30 centimeters away from the testing screen, and while horizontal viewing angle was restricted by the booth layout, vertical viewing angle was dependent on participant height. Commercially available auditory brown noise was loaded onto an mp3 player with ear plugs for use during the tactile-visual texture discrimination task. All participants first read an information sheet and signed the consent form before filing out a demographics form. Participants were tested in a quiet room with consistent lighting conditions, and asked which was their dominant hand (in the case of ambidextrous, right hand was considered the 'dominant' hand). The order of tasks consisted of the consistency test, tactile orientation task, tactile-vision texture task, tactile texture task and finally the consistency retest. The tasks took approximately one hour to complete and synaesthetes did an additional twenty minute photism-illustration task afterwards.

3.4.3.1. Touch-colour consistency test

The touch-colour consistency task used thirty novel items that were comprised of materials that lacked an implied real-world colour (e.g. wood-brown). The objects varied in weight, size, shape, texture and softness in a non-systematic manner (see fig. 3.2) to eliminate the creation of rule sets that may aid consistency for non-synaesthetes (Simner & Ludwig, 2012). A colour picker was created based on the Simner and Ludwig (2012) study, featuring a rotating colour wheel (to avoid spatial co-ordinates being used as a proxy for colour consistency), a brightness bar and a large colour preview window. Instructions were given on screen. No numeric colour information was provided to participants (i.e. RGB or HSL values), to avoid these being used to aid consistency. All aspects of the colour choice need to be actively selected by the participant to progress to the next stage (to avoid apathy resulting in inadvertent consistency). Consistent lighting conditions were provided in the testing environment to avoid inter-trial distortions in perceptual colour space.

Participants were sat down in the testing booth with the colour picker program running with instructions on screen. Participants were told they would be presented a series of objects out of sight in their non-dominant hand, and that their task was to pick the colour that best approximated either 'how the object feels in your hand' (for non-synaesthetes) or 'that best reflects your synaesthetic colour experience' (for synaesthetes). Objects were presented in a pseudo-random order and placed into their hand, palm-up for active exploration. The palm-up orientation of the participant's hand allowed the participant to gather weight information alongside other tactile attributes. Knowing that participants were able to gauge weight allowed a weight-luminance correlation to be conducted (see section 3.5.2). The ordering was determined through a single randomised order for the objects' starting points in the first presentation, and another randomised order for their starting point in the second presentation. This first and second ordering was the same across all participants. This was done to facilitate the repeated presentation of objects to the participant in a simple to understand and timely fashion across the entire task. The purpose was to prevent presentation order errors made by the experimenter while keeping a steady pace for participants in their colour selections that might be introduced in a fully random presentation.

Figure 3.2. All thirty items used in the touch-colour consistency task. Items were picked in a manner to avoid implied real world colours and varied in a non-systematic manner across tactile dimensions to avoid participants utilising rulesets which might enhance their touch-colour consistency.

For the retest, participants were re-briefed on how they would select their colours both psychologically and on the computer. When they were ready they were subsequently presented with 30 items to pair with colours in a different pseudo-random order. The retesting within an hour was chosen as consistency across non-synaesthetes and synaesthetes within short timeframes have been found to be comparable to longer retest periods (Ward et al., 2006).

3.4.3.2. Tactile orientation and texture tasks

The tactile orientation and texture tasks used commercially available JVP spatial grating domes. These have been found to provide a more reliable threshold than alternative measures of tactile acuity (Craig, 1999; Johnson & Phillips, 1981). These are small plastic domes spanning 25mm in diameter with a square wave cut into them to create a series of raised ridges and gaps that are pressed against the skin to create evenly spaced distributions of pressure. The spacing of the ridge and subsequent gap are always the same as one another on a given dome, however there are multiple domes with different spacings. There are several domes and these consist of the following spatial distances for the ridge and gap sizes: 0.35, 0.5, 0.75, 1, 1.25, 1.5, 2 and 3mm.

For the tactile orientation task, participants were told 'raised ridges' would be applied to their fingertip and their task would be to respond with the direction the ridges are orientated on the fingertip. Participants were allowed to visually familiarise themselves with the domes before a quick practice on the tactile sensation of the domes orientated along and then across the fingertip. For the task, when their hand was visually obscured and the dome was applied, participants were asked whether the dome was orientated along or across the axis of the fingertip verbally. Twenty trials were given per dome spacing and the dome progression (either progressing onto more difficult narrower domes or easier wider dome spacings) depended on the correct response rate, 15 correct stopped the task, <15 correct proceeded to the next widest ridge spacing and >15 correct proceeded to the next narrowest ridge spacing. The reasoning for this is that the final score for the tactile acquity of the participants is based on working out what dome spacing participants would be expected to get 75% of the answers correct (this is the middle point between chance and perfect performance) where a lower grating width indicating higher tactile acuity. In order to work out this 75% grating spacing the following formula was used:

$$g75 = g_{low} + (((0.75 - p_{low}) / (p_{high} - p_{low}))*(g_{high} - g_{low}))$$

g	grating spacing
p	trials correct / number of trials
high	the grating spacing or probability of correct response on the lowest grating spacing on which the participant responded correctly more than 75% of the time
low	the grating spacing or probability of correct response on the highest grating spacing on which the participant responded correctly less than 75% of the time
*g*75	the hypothetical grating spacing on which the subject would have scored 75% had it been present

No feedback on performance was given and the order of presentation was pseudo-randomised. The domes were applied manually until a response was given as performance is relatively unaffected by the force of application (Johnson & Phillips, 1981). To reduce time taken for the experiment, the 1mm dome was chosen as the starting point as it forms the middle ground between typical tactile acuity and enhanced mirror-touch synaesthetic acuity (Banissy et al., 2009).

For the tactile texture task, participants were shown two JVP domes of 1.5mm and 1.25mm width. These were described as the wide and narrow domes respectively, and that the participant's task was to verbally identify which of the two is administered to them tactually across their non-dominant hand's index fingertip. The comparison conditions included twenty trials of 1.5mm vs 1.25mm, 1.25mm vs 1mm and 1mm vs 0.75mm. The order of wide and narrow presentation was done in a pseudo-random order. Stimulation was manually applied until a response was given and a higher percentage correct denotes a higher tactile acuity.

3.4.3.3. Tactile-visual texture discrimination task

The tactile-visual discrimination task used commercially available aluminum oxide sandpaper for variations of tactile texture. The GRIT values (reflecting particles per square inch, with smaller numbers of particles denoting larger particles and a rougher texture) of 40 (the roughest), 60, 80, 120, 180 and 240 (the smoothest) were used. For the visual presentation of texture, photographs of each type of sandpaper were taken under consistent lighting conditions; the images were then converted to greyscale and several versions of these images were created through rotating the image to avoid identical visual repetition or cues to the participant. The pictures were analysed in terms of their mean luminance using GIMP image manipulation software. Analysis revealed that the pictures had minor random variations in mean luminance across all of the GRIT values. Since we did not want participants to mistake variations in average luminance as a cue to which visual texture they were seeing, a standardisation process was applied. All images were edited to have equal luminance ratings of 140 out of 200. Besides the overall luminance of an image there is also the amount of variation present between the lighter and darker elements of an image. When it comes to the variation of brightest and darkest individual parts of an image, since rougher GRIT values had larger particles, these naturally created larger and darker shadows than the smoother GRIT values. This variation in lightest and darkest elements of an image can be analysed in terms of the average standard deviation a pixel has from the mean luminance of the image as a whole. The roughest GRIT value had a standard deviation of 33 and the smoothest had a standard deviation of 15. This confirms that rougher GRIT values have darker and more luminance elements of an image (while at the same time the mean luminance is equal) relative to the smoother GRIT values that only show subtle variability from the mean luminance. Using GRIT values as a guide we adjusted the standard deviation from the mean luminance value to be linear between these points. This gives us the following standard deviations for each image: 40 GRIT (SD = 33), 60 GRIT (SD = 31), 80 GRIT (SD = 29), 120 GRIT (SD = 25), 180 GRIT (SD = 21), 240 GRIT (SD = 15). As such all pictures are equally luminant, however variations between the darker and lighter parts of an image are larger for lower GRIT ratings which visually indicates that the particles are larger and hence rougher than smoother GRIT values.

Participants were allowed to freely view pictures of the sandpaper for a minute that would later be used in the experiment. They were then explained the purpose of the experiment was would be to indicate whether sandpaper that is tactually felt is the same as or different to, the sandpaper that is presented visually on the computer. During the task, auditory cues as to perceived roughness were masked via the presentation of auditory brown noise on an mp3 player (Guest, Catmur, Lloyd & Spence, 2002). The participant's dominant hand would tactually explore the sandpaper presented for five seconds before the visual sandpaper is presented, participants would then indicate whether the tactile and visual sandpapers were the same or different with the ‘S’ and ‘D’ keys respectively with their non-dominant hand. The visual presentation and recording of responses was done using e-prime 2.0 professional. The program featured six familiarisation, six practice (with feedback) and sixty trial phases (without feedback).

3.4.3.4. Photism-illustration task

In order to explore the psychophysics of how tactile stimulation relates to the visual photisms of the synaesthetes, eighteen photism-drawing tasks were devised. Synaesthetes were presented with answer sheets and commercially available drawing materials ranging across the colour spectrum. Synaesthetes were instructed to have their hand, body and head at specific orientations for each condition. They were then told that the experimenter would administer tactile stimulation three times on their palm while they had their eyes closed. The synaesthetes were then asked draw the photisms experienced from this stimulation. Any motion of the photism would be indicated by descriptions or arrows. The specific orientations of the hand and types of tactile stimulation are detailed in the results section.

3.5. Results

3.5.1. Touch-colour consistency task.

Each colour selected by participants was transformed from their original RGB values into CIE LUV space (assuming a standard colourimetric oberserver of a 2° visual angle and D65 illumination – see section 1.4.9). The conversion to CIE LUV is done to use a colour space more closely based on human colour perception. In CIE LUV space, Euclidian distances between two points reflect perceived distances to the human observer, consisting of three dimensions based on luminance as well as red-green and blue-yellow colour opponency. This colour space has been found to be the most accurate in distinguishing synaesthetes from non-synaesthetes for grapheme-colour consistency tests, even without the use of colour-corrected monitors (Rothen, Seth, Witzel & Ward, 2013). Distances between the colours selected for each object created an error score for each object, with the average across all objects creating the participant's final error score with lower values denoting higher colour consistency.

The consistency scores for the control group had a mean distance of 122.18 (SD = 38.80) CIE LUV units for colours chosen for the same object between the initial test and re-test tasks. This control group was compared to individual synaesthetes using Crawford's modified t-test for single case data (Crawford & Garthwaite, 2002). A one-tailed test was run for all synaesthetes which revealed that only ES was significantly more consistent than controls with an average of 51.54 CIE LUV units of error per object (t = -1.79, p = .042). MLS and CS were not significantly more consistent than controls with CIE LUV error scores of 74.86 and 66.52 units per object, however they did tend in this direction with t = -1.20, p = .120 and t = -1.41, p = .084 respectively. The most consistent synaesthete, ES, had the most consistent colour pairings across all participants. Overall the tactile-vision synaesthetes were significantly more consistent as a group than controls using a one-tailed Mann-Whitney U Test, U (34) = 5, Z = -2.52, p = .006 (one-tailed), d = 1.49 (see fig. 3.3) with a mean of 64.31 (SD = 11.81) CIE LUV units error over the retest periods (see fig. 3.3).

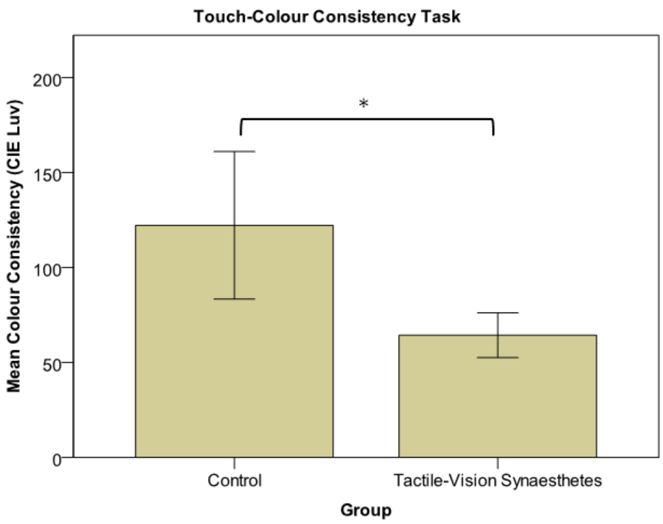


Figure 3.3. Touch-colour consistency task as measured in CIE LUV colour units for the control and synaesthetic groups. Lower values indicate greater consistency between initial test and re-testing periods. Tactile-vision synaesthetes had significantly more consistent touch-colour pairings than controls on this task. Error bars show 1 standard deviation from the mean. Key: * = p < .05.

3.5.2. Weight-luminance correlations.

There are a variety of correspondences that participants can utilise in matching colours to tactile objects. One previously identified correspondence is that of weight-luminance, where heavier objects yield darker colours and lighter objects more luminant colours (Walker, Francis & Walker, 2010). Due to the presentation of the stimuli into the participants' up-turned palms, this should aid in the processing of weight information. As a result, this was an opportune time to further explore this correspondence and whether it has any influence on synaesthetic photisms.

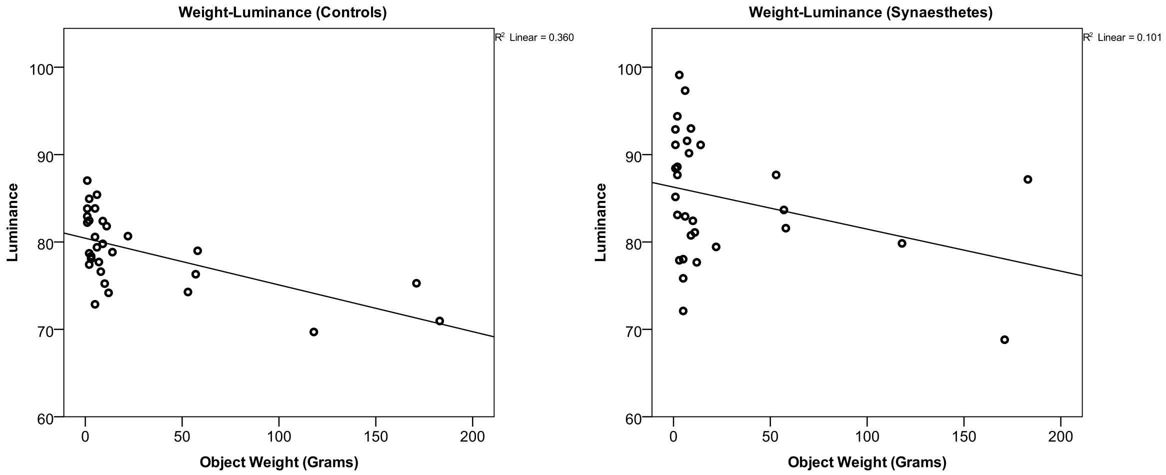


Figure 3.4. Scatterplot showing the correlation between object weight and luminance of colours chosen during touch-colour consistency task for controls (left) and tactile-vision synaesthetes (right).

A Pearson's *r* correlation was performed between the object weight (grams) and the average luminance values for each object across test-retest periods and across participants within each group. For the non-synaesthetic group, object weight and luminance were negatively correlated, so that increases in weight corresponded to decreases in luminance, $r\ (30) = -.60$, $p < .001$ (one-tailed) as can be seen in figure 3.4. A similar but weaker correlation was found for synaesthetes with $r\ (30) = -.32$, $p = .044$ (one-tailed), who experienced darker visual photisms to heavier objects. Both of the correlations were found to be in the same direction as previous studies on weight-luminance correspondences (Walker, Francis & Walker, 2010; Ward et al., 2008).

3.5.3. Discrimination tasks

During the behavioural discrimination tasks, synaesthete ES reported that the presentation of JVP domes to the fingertips resulted in substantial illusory heat being felt at that location. The illusory heat resulted in several stops during the tasks to recover with ES reporting it as masking normal tactile sensations. As a result, ES' scores are excluded from the tactile texture and orientation tasks. There were no previous examples of this touch-temperature sensation reported by the participant, so it is likely that the distribution of ridges of the JVP domes induced this sensation.

3.5.3.1. Tactile texture task

For the tactile texture task using the JVP domes (see fig. 3.5), a Mann-Whitney U test indicated that tactile-vision synaesthetic group (Mean = 79.00, SD = 8.48) achieved a significantly higher score than controls (Mean = 64.23, SD = 11.29), with U (33) = 6.50, Z = -1.85, $p = .032$ (one tailed), *d* = 1.31. This indicates that the synaesthetes' ability to discriminate variations in spatial texture on the skin is enhanced relative to controls.

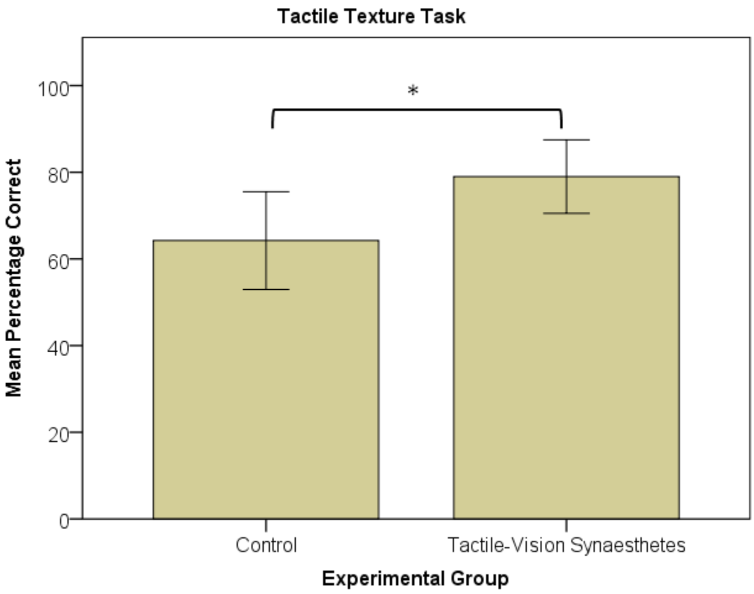


Figure 3.5. Bar chart showing the mean percentage correct responses for the tactile texture task for the control and synaesthetic groups. Tactile-vision synaesthetes had significantly higher tactile texture discrimination relative to controls on this task. Error bars show 1 standard deviation from the mean. Key: * = $p < .05$.

3.5.3.2. Tactile orientation task

For the tactile gratings orientation task (see fig. 3.6), a Mann-Whitney U test indicated that there was no significant difference between the tactile-vision synaesthetic group (Mean = 1.75, SD = 1.06) and controls (Mean = 1.95, SD = 0.65), with U (33) = 22.50, Z = -0.642, p = .261 (one tailed). This indicates that the tactile-vision synaesthesia group does not have an advantage in tactile acuity relative to controls. There was considerable variability in scores within the tactile-vision synaesthesia group with MLS showing the highest tactile acuity of all participants. However MLS also has mirror-touch synaesthesia which has been previously demonstrated to result in higher tactile acuity on the tactile orientation task (Banissy et al., 2009).

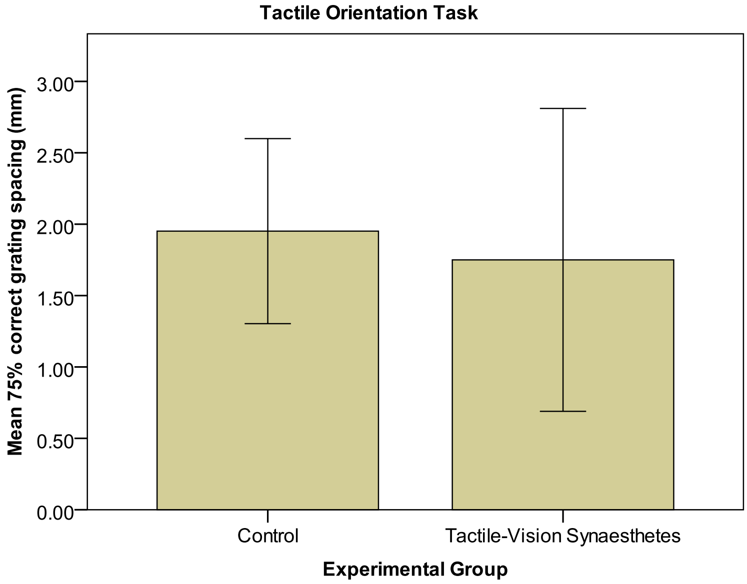


Figure 3.6. Bar chart showing the hypothetical grating width that would achieve a 75% correct score by participants. Lower scores indicate better tactile acuity. The control and synaesthetic groups did not significantly differ on this measure. Error bars show 1 standard deviation from the mean.

3.5.3.3. Tactile-visual texture task

For the tactile-visual texture comparison task, a Mann-Whitney U test indicated that there was no significant intergroup difference between tactile-vision synaesthetes (Mean = 71.67, SD = 11.06) and controls (Mean = 69.16, SD = 5.69) in terms of mean percentage correct (see fig. 3.7), U (30) = 29.50, Z = -0.76, *p* = .223 (one tailed), *d* = 0.44. Similar to the orientation task, there was a large amount of variability within the synaesthesia group, with MLS having the highest score of all participants. Important to consider is that MLS is also the only synaesthete with both mirror-touch and grapheme-colour synaesthesia, which have been found to be associated with higher tactile-acuity and increased neural response patterns to high-spatial frequency visual content, which may aid in solving this task through enhanced discrimination of both relevant modalities (Banissy et al., 2009; Barnett et al., 2008b). Despite the tactile-visual synaesthetes' better discrimination of tactile spatial frequency and most of the synaesthetes also having grapheme-colour synaesthesia, this did not lend itself to an advantage for all tactile-visual synaesthetes.

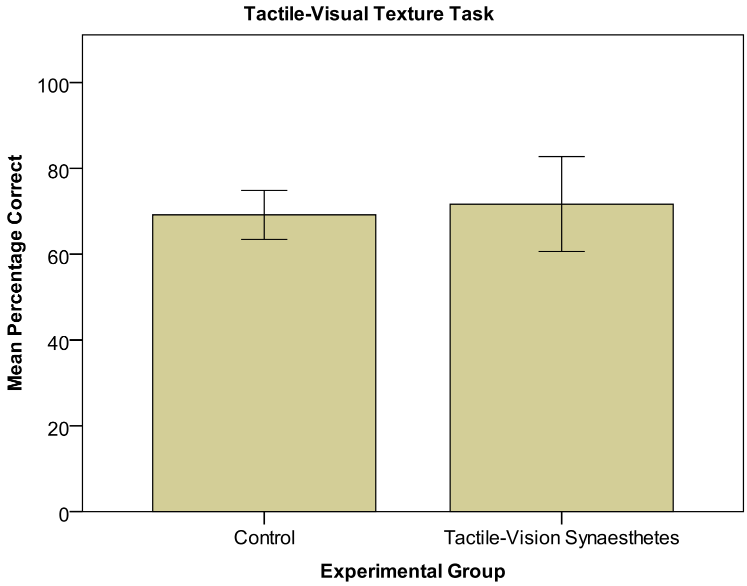
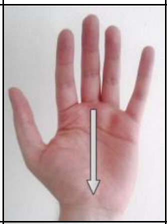
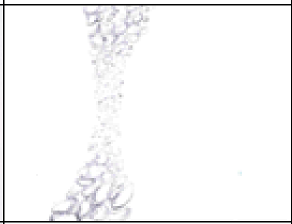
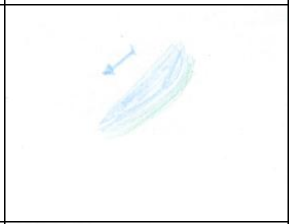
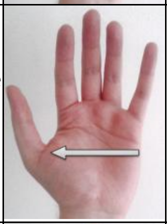
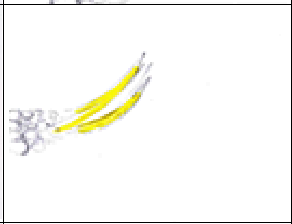
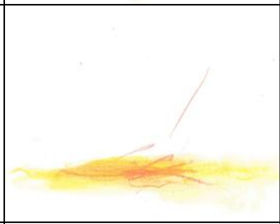
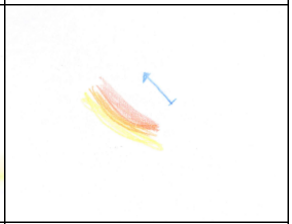
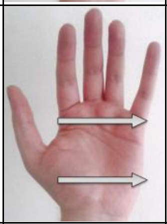
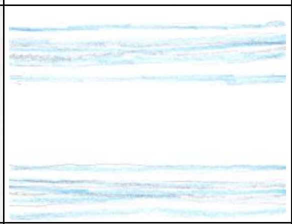
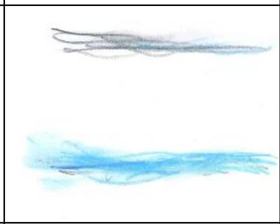
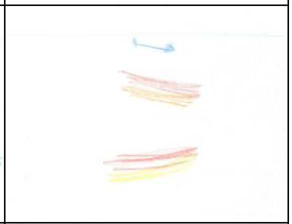
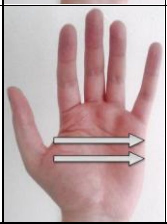
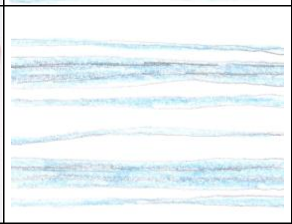
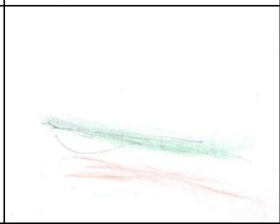
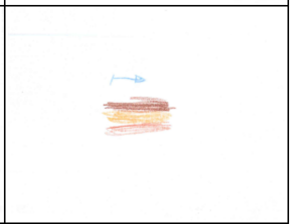
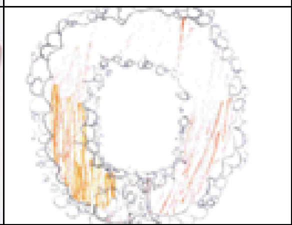
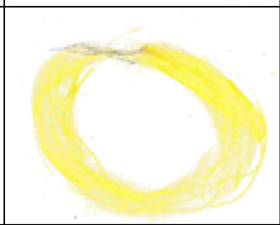
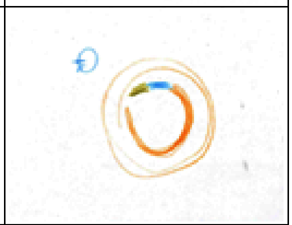
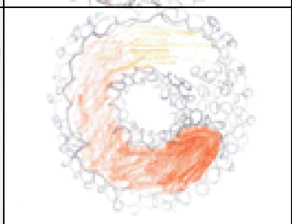
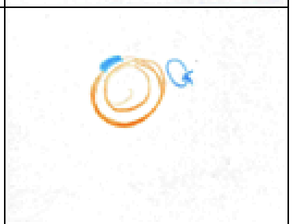
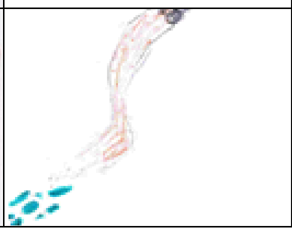
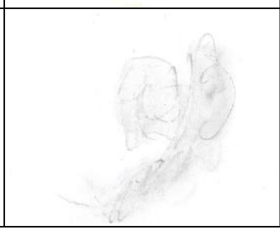
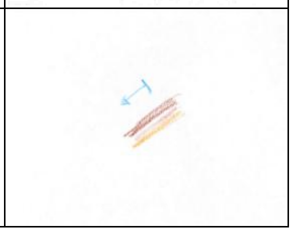


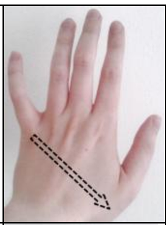
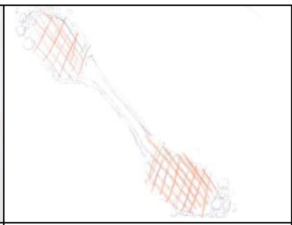
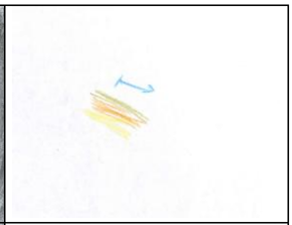
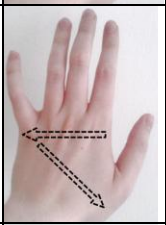
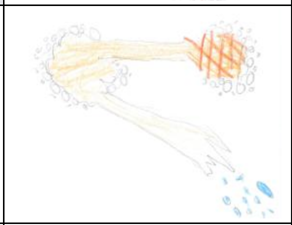
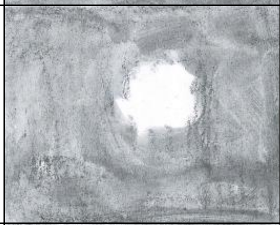
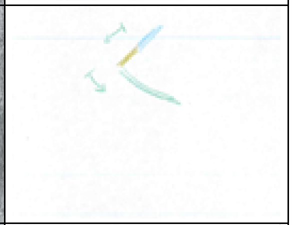
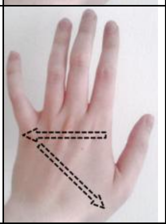
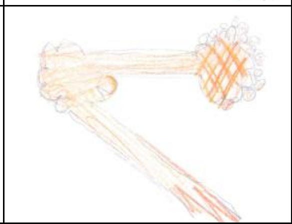
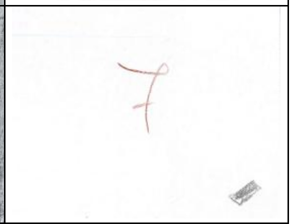
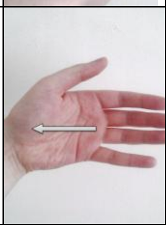
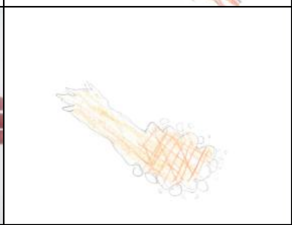
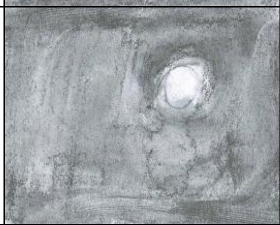
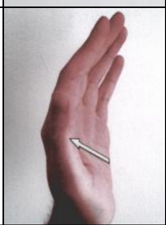
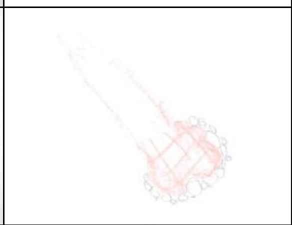
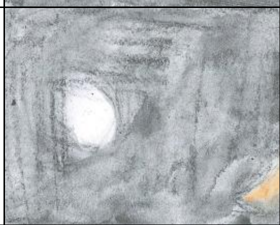
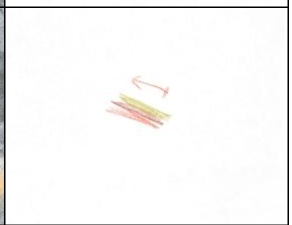
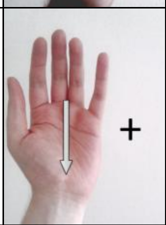
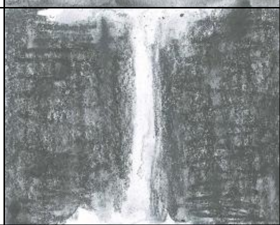
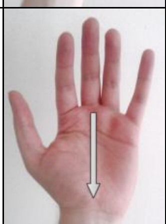
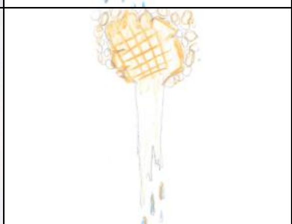
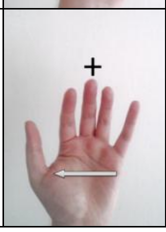
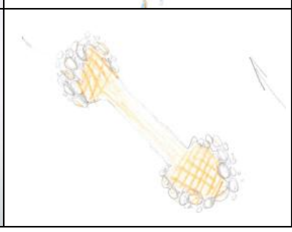
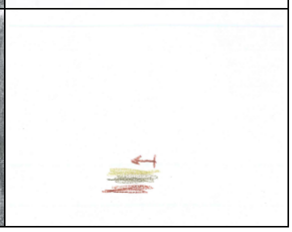
Figure 3.7. Bar chart showing the mean tactile-visual texture task scores for the control and synaesthetic groups. The groups did not significantly differ in terms of percentage correct identification of tactile and visual textures matching. Error bars show 1 standard deviation from the mean.

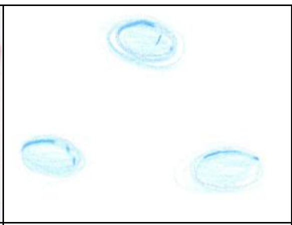
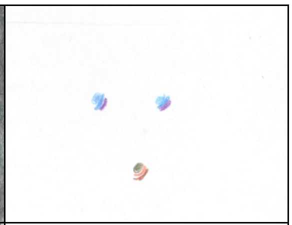
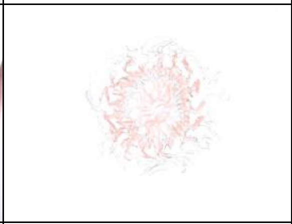
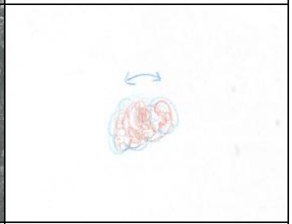
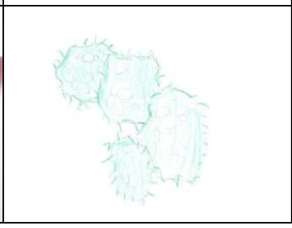
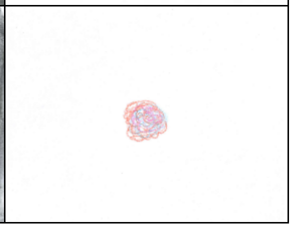
3.5.3.4. Photism illustration task

In order to better understand how the tactile system influences the visual system in tactile-vision synaesthesia, a series of variations in touch was given to the each synaesthetes' left hand. Participants were tasked with drawing their visual photism in response to each of these tactile stimulations. The tactile stimulation was designed so that contrasting different stimulations allows an examination of how the photism varies in response to spatial changes in pressure, motion and perspective. The default hand position for participants involves their elbow resting on the table, open palm facing themselves at head level with fingers pointing upwards. Each touch is delivered three times while the participants had their eyes closed. Participants were asked to consider the centre of the drawing box to reflect the centre of their visual field. The full list of variations in stimulation and the photisms drawn in response by synaesthetes can be seen in table 3.3.

Table 3.3. Photism illustrations in response to tactile stimulation.

	Stimulation	Illustration	MLS	CS	ES
1	Default hand position, downward stroke				
2	Default hand position, leftward stroke				
3	Default hand position, rightward stroke, with wide gap				
4	Default hand position, rightward stroke, with narrow gap				
5	Default hand position, clockwise circular stroke				
6	Default hand position, anti-clockwise circular stroke				
7	Default hand position, diagonal stroke, little finger to thumb				

8	Palm facing away, diagonal stroke, little finger to thumb				
9	Palm facing away, draw a '7' on the palm				
10	Palm facing away, draw a '7' on the palm, ask "what number is it?"				
11	Fingers pointing right, stroke leftward				
12	Palm facing right, stroke from furthest to nearest point				
13	Default hand position, face looking right, stroke downward				
14	Default hand position, body facing right, stroke downward				
15	Default hand position, hand at chest level, stroke leftward				

16	Default hand position, pressure on three points				
17	Default hand position, ball rotated in palm				
18	Default hand position, ball held between fingertips				

Key: Solid line = stimulation to near side; dotted line = stimulation to furthest side. Arrows = direction of tactile stimulation. Fixation cross = orientation of head.

3.5.3.4.1. Spatial stimulation

The first variable is how changes in spatial stimulation affect synaesthetic photisms, with strokes along the vertical (1) and horizontal (2) axes producing photisms that are rotated 90 degrees with respect to each other for all participants. The depth axis (12) is also experienced for MLS with descriptions of their photisms going "towards" them, however CS shows no obvious influence of this, and ES' photisms are more in line with the previous horizontal stimulation. The distance between touched regions (3 & 4) show a literal representation of spatial distance, with wider tactile gaps producing wider visual gaps while narrower gaps reduce this distance (MLS & CS) or eliminating it entirely (ES). Other variations in spatial pressure also seem to have a closely knit relationship to visuospatial representations, with circular strokes (5, 6) producing circular photisms and pressure on three points on their hand produce 'dots' that appear to represent the allocation of pressure (16) although not always accurately (CS). Finally the representation of objects was tackled in two ways, with the passive receiving of a circular object into their palm (17) or holding the object in their fingertips (18). Synaesthetes uniformly had circular photisms to the passive receiving of the object, but are split on how holding the object is visually represented. For holding an object in their fingertips, MLS experienced photisms with multiple points, possibly representing the contact points between their fingers and the object, quite differently CS and ES both have circular representations closer to the passive representation of the object.

3.5.3.4.2. Perspective

Varying the relationship between the individuals' perspective and the spatial location of tactile stimulation has been found to modulate photisms in acquired synaesthesia (Armel & Ramachandran, 1999), this section will explore if this is also true of developmental synaesthesia. Situations in which the proximal stimulation stayed the same (i.e. identical stroking patterns) but changes to the positioning of the hand in relation to the face are examined first. When the synaesthetes had a diagonal line (little finger to thumb) stroked into their palm while it was facing them (7), MLS & ES reported photisms that moved from top left to bottom right, however when the same stimulation is given to the palm facing away (8), the photism 'flips' for these synaesthetes so that it remains consistent with their visual perspective. When the hand is moved in relation to the face, MLS and CS have their photisms remain central while ES reports that their photisms move horizontally (13) or vertically (15) with their hand, with ES also reporting that they occur outside of her visual space. Changing the body's position while keeping the head position constant does not appear to influence photisms for any synaesthetes (14).

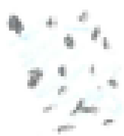
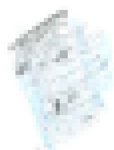
3.5.3.4.3. Motion

Synaesthetes were asked to report the direction of any visual motion that occurred from the touch stimulation. CS did not explicitly report any visual motion from the stimulation for any stimulus. Clockwise (5) and anti-clockwise (6) tactile stimulation produced clockwise and anti-clockwise motion in MLS and ES' synaesthetic photisms respectively. Likewise, horizontal (2), vertical (1) and towards (12) stroking motions produced congruent motions in MLS and ES as well.

3.5.3.4.4. Texture

ES reported that exploring textures in particular had a strong influence on their photisms, so stimuli from the consistency test were brought back for her to explore and report her photisms to. The results can be seen in table 3.4.

Table 3.4. Photisms to active exploration of different textures by ES

Rough sandpaper (40 grit)	Sandpaper (240 grit)	Fabric	Smooth ceramic tile
			

The rougher grainier textures appear to have more isolated 'specks' in the photisms of ES, similar to the previously delivered three points of pressure (16). As the textures become smoother the photisms become more integrated and uniform in their visual texture, and colour saturation appears to increase with smoothness, similar to correspondences in non-synaesthetic children and adolescents (Ludwig & Simner, 2013) but opposite to a previously investigated synaesthete, for whom roughness related to increased saturation (Simner & Ludwig, 2012).

There are many interesting commonalities in how synaesthetic photisms change in response to variations of touch. The first is that the spatial organization of touch relative to their visual perspective appears to be reflected in the spatial distribution of colour in their photisms. Despite this, varying the location of the hand had mixed effects with the photisms following the hand position for ES, or remaining central for MLS and CS. The orientation, spatial distribution and motion of strokes appear to be accurately mapped in most of the synaesthetes' photisms (however CS did not report motion). Orientation and motion processing involve shared sensory areas between tactile and visual processing (Blake et al., 2004; Hagen, et al., 2002; Zangaladze et al., 1999; Zhang et al., 2005). The spatial distribution of tactile stimulation appears to have a close relationship with the distances between colours in the synaesthete's photisms. The ability to accurately interpret distances between points on the skin appears to be enhanced in tactile-vision synaesthesia from earlier experiments using the JVP texture task. This brings two previous bodies of work into consideration. Firstly it could mean the use of the RAG, an area used in interpreting distances between skin locations (Sathian et al., 1997; Spitoni et al., 2013; Zhang et al., 2005). Secondly, it could be the result of early visual area recruitment which has been incorporated for interpreting tactile distances in work on visually deprived individuals (Merabet et al., 2004, 2007, 2008; Sathian & Stilla, 2010). Other potential avenues for future investigation include research into shared external spatial representations for both the tactile and visual modalities (Spence, Pavani & Driver, 2000). For instance, the processes that interpret external tactile spatial cues affect early visual processing (Eimer & Driver, 2000) or how multi-sensory representations of external stimuli both involve the intraparietal sulcus (Macaluso & Driver, 2001). Together these may provide clues as to how such a tight spatial mapping is achieved between the inducing touch and concurrent photisms in synaesthesia.

3.6. Discussion

The present experiment represents the largest questionnaire and behavioural study to date on how both touch and visual photisms are influenced through the presence of tactile-vision synaesthesia. In a sample of 21 synaesthetes, respondents noted that the most common inducers of synaesthesia were emotional human touch, itchiness and sexual experiences. Variations in textural quality appeared to vary the content of visual photisms the most, which typically consisted of colour, motion and texture. Some synaesthetes do not report consistency or that inanimate objects create photisms which may make some unsuitable for consistency testing. In our consistency testing, one synaesthete reached significance, while others reporting consistent photisms had a tendency in that direction. In our behavioural testing, synaesthetes have a significantly superior ability to discriminate distances on the skin, but did not show superior orientation or tactile-visual integration abilities. These findings indicate either superior RAG involvement or the recruitment of additional visual areas to aid in discrimination. Finally, the photisms themselves appear to contain a close tactile to visual spatial mapping and depend on their visual perspective rather than sorely on the skin location of tactile stimulation.

3.6.1. Consistency testing

Of the three synaesthetes that underwent consistency testing, only one reached significantly higher levels of consistency than controls, while the final two tended in that direction. As a group, these three tactile-vision synaesthetes are more consistent than controls. This data suggests that consistency may indeed be a suitable measure of genuineness for some tactile-vision synaesthetes. During consistency testing, controls use many tactile-visual correspondences which aid their consistency, making the test less sensitive for identifying synaesthetes (Ludwig & Simner, 2013; Ward et al., 2008). Another factor potentially reducing sensitivity in the present experiment is the time period of the test-retesting being an hour unlike Simner and Ludwig's (2012) case study, which tested a synaesthete over four months and sixteen days for controls. While shortened periods of retesting may be suitable for other forms of synaesthesia (Ward, Huckstep & Tsakanikos, 2006) this may provide too much of an aid to controls in maintaining their touch-colour pairings. Further sensitivity could be added in several ways; firstly objects that show high consistency for controls could be omitted from further testing as they would not be sensitive to identifying synaesthetes. In addition future studies could either focus on tactile dimensions for which only synaesthetes have consistent-mappings or by introducing distracters for controls that do not affect synaesthetes' photisms. For example, since texture is the strongest reported modulator of colour for synaesthetes applying the same texture to objects varying in dimensions that do not influence synaesthetic photisms, might make texture-colour consistency a stronger measure of genuineness. However this still leaves the problem of testing synaesthetes that do not have photisms to inanimate objects. A remaining commonly reported inducer is that of human touch which may also be suitable for consistency testing, variations in orientation and pressure seemed to influence the colour of the photisms reported in the photism-illustration task, in addition, human touch could have variations in texture through additional materials at the point of contact. This also has the added benefit of being passively received, since variations in exploration strategies by users may add to inconsistent tactile sensations and potentially more inconsistent colour selections as a result.

3.6.2. Weight-luminance

Previous investigations into touch-colour links for controls and synaesthetes found multiple associations between softness, roundness and smoothness with luminance and saturation (Ludwig & Simner, 2013; Simner & Ludwig, 2012). This study adds further evidence of a negative weight-luminance correspondence found in non-synaesthetes (Walker et al., 2010), and furthermore that this has an impact on synaesthetic photisms. Whether weight is the primary cause of the luminance association is unclear however, as previous negative pressure-luminance associations have been observed in controls and synaesthetes, and could also explain the present findings (Ward et al., 2008). One way these explanations could be disentangled is through changing the surface area that an object makes contact with the skin, this way the same weight can be felt as higher or lower pressures, or pressure can be kept constant while overall weight changes.

3.6.3. Behavioural tasks and neural processes

Behavioural testing indicated that the tactile-vision synaesthetes had an enhanced discrimination of spatial texture but not orientation or tactile-visual integration relative to controls. The neural correlates of discrimination on the texture task indicate that only the right angular gyrus (RAG) is more active then when compared to the orientation task (Zhang et al., 2005). The RAG is functionally involved in discriminating distances between locations on the skin (Spitoni et al., 2013). Lesions to the RAG can result in deficits for body image (location of limbs) but not body schema (number and type of limbs) in reporting tactile localisation (Anema, Kessels, de Haan, Kappelle & Leijten, 2008). Similarly it has been suggested that the RAG is crucial in transforming co-ordinates from motor reference frames to visual reference frames (Muggleton, Cowey & Walsh, 2008). As such, if the improved discrimination is primarily due to the RAG, we might also expect superior tactile localisation on the body as well. Further underlining the use of the RAG for tactile-visual synaesthesia is the tight mapping between the location of human touch and location of colours seen in the photism-illustration task. It might be expected that the visual concurrents may be able to assist in spatial discrimination similar to the enhanced temporal discrimination seen from auditory concurrents (Saenz & Koch, 2008). However during discrimination tasks CS and ES denied any influence of the task on their photisms, with MLS noting that only the orientation task influenced the orientation of her photisms (other tasks had no effect). MLS is unique here in that the presence of mirror-touch synaesthesia is already known to raise tactile sensitivity (Banissy et al., 2009), this in combination of touch-colour synaesthesia might mean that given a certain level of tactile sensitivity orientation information specifically from inanimate objects might be enough to have the POC influence subsequent photisms, marking it as a viable route in her case of touch-colour synaesthesia. Whether this is applicable to the other touch-colour synaesthetes might only be evident if their tactile sensitivity were raised, such as was previously seen using tDCS (Ragaert et al., 2008). Despite using the RAG, the texture task appears not to automatically induce informative photisms to the task, so more powerful inducers like human touch may only be able to 'break through' into consciousness. An alternative possibility is the recruitment of new processes well suited to spatial discrimination as seen with V1 recruitment for tactile texture discriminations in visually deprived populations (Merabet et al., 2004, 2007, 2008). In support of this is that tasks that already use visual processing for controls are not enhanced for synaesthetes, but the texture task that does not normally have visual processing *is* enhanced.

In examining other regions used in the JVP texture task, unfortunately there is no neuro-imaging study directly showing a comparison between the JVP texture task and a baseline of no tactile stimulation, however Zhang et al.'s (2005) data features texture, orientation and no-stimulation conditions. From this it is possible to gauge which regions are used in the JVP texture task relative to baseline by identifying regions that are significantly related to the ‘orientation > no tactile stimulation’ condition and are no longer significant in the ‘orientation > texture’ condition. Regions that meet these criteria include: Left post-central sulcus, left parietal operculum, bilateral frontal operculums, inferior frontal gyrus, right cerebellum and dorsal pre-motor regions. The primary and secondary somatosensory cortices are within this post-hoc comparison (Keysers et al., 2010) and are likely to be used in both the JVP texture and orientation tasks. Supporting this suggestion is the knowledge that other forms of tactile texture activate both the primary and secondary somatosensory cortices (Burton, Macleod, Videen & Raichle, 1997; Stilla & Sathian, 2008) and that inhibition of the primary somatosensory cortex using TMS can revert texture and orientation discrimination performance to chance levels (Zangaladze et al., 1999). Conversely lowering the neural firing threshold for S1 through rTMS or tDCS can enhance discrimination for orientation and spatial texture tasks (Ragert et al., 2008; Tegenthoff et al., 2005). As such tactile-vision synaesthetes are unlikely to show functional changes to S1 and S2 relative to controls, otherwise we would see enhancements to discrimination for *both* the texture and orientation tasks, not just texture.

In a questionnaire given to 21 tactile-vision synaesthetes, the most commonly reported inducers include emotional human touch, itchiness and sexual experiences, in addition to this, variations in texture are reported to influence the subsequent photism. Common to all of these points is the involvement of the insula (Deshpande et al., 2008; Herde et al., 2007; Hole et al., 2012; Komisaruk et al., 2004; Morrison et al., 2011; Olausson et al., 2002; Valet et al., 2008). The insula also is involved in temperature and pain processing (McGlone, Vallbo, Olausson, Loken & Wessberg, 2007; Singer, Critchley & Preuschoff, 2009), which are inducers for certain forms of synaesthesia which have a high co-morbidity with tactile-vision synaesthesia (Novich et al., 2011). As such the insula is likely to be a strong modulator of touch for inducing tactile-vision synaesthesia.

If the insula is indeed involved in the specific mapping of touch-colour synaesthesia, we might also expect similar mappings between visual and emotional dimensions as might be seen in 'emotion-colour' synaesthesia. Currently emotion-colour synaesthesia is known to more reliably invoke colour for words with strong emotional content and that the type of emotion (positive / negative) also has some influence on the specific colours selected (Ward, 2004). The colours chosen for positive words tended towards highly saturated colours such as yellow, while for negative words black was overwhelmingly selected. These results follow trends between emotions and colours for non-synaesthetes (Collier, 1996; D'Andrade & Egan, 1974; Palmer et al., 2013). The preference order for emotions and colours independent of one another are closely related to the association between emotions and colours, especially for children (Terwogt & Hoeksma, 1995), since synaesthesia is a developmental condition, the prevalence of these links in childhood may become more firmly rooted as synaesthesia manifests, maintaining the association further into adulthood (Ludwig & Simner, 2013; Simner & Ludwig, 2012). Specific colours have been reported to co-inside with the level of familiarity to the person triggering the synaesthesia and these colours become fixed when the personality of the inducing person is firmly known (Collins, 1929; Cytowic, 1989; Ramachandran, Miller, Livingstone & Brang, 2012; Riggs & Karwoski, 1934). In some other individuals, thinking of an emotion either as a concept or as a word can elicit congruently emotional colours (Raines, 1909) while unexpected swathes of emotion can tint their visual perspective (Cutforth, 1925). Furthermore in breaking down a link between 'touch to emotion to colour' are cases of tactile-emotion synaesthesia, where specific textures (e.g. silk) evoke specific emotional states such as relief

(Ramachandran & Brang, 2008). It was speculated by Ramachandran and Brang (2008) that this process might involve additional connections between the insula and somatosensory cortices. If emotion-colour forms the basis of touch-colour associations, we would expect that touch-emotion, emotion-colour and touch-colour associations would all be congruent with one another, and so positive variations of touch would be more likely to evoke yellowish colours, while negative sensations are more likely to be black. Whether emotion is the true elicitor of the colour sensations or whether emotion plays a guiding role in the formation of touch-colour synaesthesia remains to be determined.

3.6.4. Conclusion

In a series of experiments, we present questionnaire and behavioural data from the largest group of tactile-vision synaesthetes gathered to date. The questionnaire data suggests common inducers for this synaesthesia which also point to specific underlying neural mechanisms such as the insula in modulating tactile sensations. The phenomenology of this synaesthesia also shows a close spatial mapping between the location of touch and vision implicating mechanisms involved in translating tactile to visual spatial reference frames such as the right angular gyrus. Further to this, tactile-vision synaesthesia is associated with enhanced spatial discrimination on the skin, another process that specifically involves the right angular gyrus. Replicating previous measures of consistency-testing yielded only one self-reported synaesthete to significantly outperform controls, with another two tending in that direction. Overall, this evidence indicates that the presence of touch-colour synaesthesia results in several perceptual and behavioural effects. Taken together, this implicates specific neural mechanisms that may form the basis for future research into a wide variety of interactions between the tactile and visual systems.

4. Colour in Sensory Substitution: Object Discrimination and Colour Replication

4.1. Abstract

Sensory substitution devices (SSDs) translate visual dimensions into patterns of sound or touch, which allow the visually-impaired to access the 'visual world.' Previous SSD research has focused on greyscale vision, however, evidence from computer-vision and low-resolution vision studies suggests that dimensions of hue and saturation may aid in the extraction and identification of objects in complex real-world environments. No studies have compared whether the information content (greyscale, colour) or presentation of information (i.e. luminance represented by pitch or loudness) is important for users of these devices. To investigate the effect of different colour representations, we used the 'Creole,' a tablet that translates a single point of colour underneath the users' fingertip into variations of auditory and tactile stimulation (for further details see section 4.3.3.1). The cross-sensory mappings were based on prior sensory substitution devices and the cross-modal correspondence literature. After 15 minutes of training, participants used the Creole to discriminate which specific fruit/vegetable they were listening to on the Creole from a selection of four options, varying in fruit/vegetable and environmental lighting (high / low) in a 4AFC task. Subsequently users attempted to reproduce 10 colours from sounds representing 10 randomly chosen colours. Participants discriminated fruit/vegetables significantly better with luminance-loudness mappings, irrespective of whether there was colour information. An analysis of errors made by participants found that luminance-loudness mappings were more accurate in establishing the overall lighting. Interestingly, in colour replication tasks, luminance-loudness and luminance-pitch approaches performed at equivalent levels, suggesting that the difference in luminance-errors during the object discrimination task was not due to user's inability to discriminate luminance, but rather to utilise this information efficiently when exploring the complex lighting variations present in natural stimuli. Finally, the colour-mappings that performed best at object discrimination, performed worst at colour replication, indicating that the best approaches for natural stimuli are not dependent on the most accurate colour knowledge. Overall, our evidence suggests that the presentation of information has the most significant impact on the user's performance, that different presentations suit different tasks and that not all correspondences work equally well in sensory substitution.

4.2. Introduction

Visual sensory substitution devices (SSDs) allow the properties of vision to be understood or re-experienced through other senses. The most researched devices convert 2D greyscale images consisting of vertical position, horizontal position and luminance into dimensions of touch or sound. For instance, the 'Brainport' SSD displays these images on the users' tongue using a combination of skin position and electrotactile intensity, while the 'vOICe' SSD converts these images into patterns of pitch, time and loudness (Bach-y-Rita, Kaczmarek, Tyler & Garcia-Lara, 1998; Meijer, 1992). When a user moves the visual sensor a feedback loop occurs, so that interactions with the environment produce predictable changes in the SSD signal. This interaction results in the transmission of the sensorimotor aspects of vision such as perspective, occlusion, and visual size, but experienced through touch or hearing (Amedi et al., 2007; Bach-y-Rita, 2004; Collignon, Lassonde, Lepore, Bastien & Veraart, 2007; Ward & Wright, 2014). With continued training, SSD experts not only learn to decode the patterns of touch and sound back into a 'visual' image, but start recruiting visual regions of the brain for SSD processing in the blind (Amedi et al., 2007; Matteau, Kupers, Ricciardi, Pietrini & Ptito, 2010; Merabet et al., 2009; Ptito, Matteau, Gjedde & Kupers, 2009). Some blind SSD experts have reported visual phenomenology such as location and luminance from these SSD signals (Ward & Meijer, 2010). Interestingly, this phenomenology can include experiencing depth, motion and colour using devices that do not directly encode this information. Colour inference in particular appears to be guided through a combination of recognition and prior visual information so while 'Christmas trees' are filled-in green, ambiguous objects such as clothing may evoke a more arbitrary colour based on best information available. The next section will explore how the direct encoding of colour in SSDs may overcome some of the ambiguities of greyscale. This is followed by an examination of different approaches to colour as well as their representations in touch and sound.

4.2.1. Utility of colour in SSDs

In tasks that require figure-ground segmentation, such as object discrimination, recognition and navigation, the contours that separate objects from their background may not always be best defined by changes in luminance. However other dimensions of colour such as hue and saturation provide additional ways to segregate an image, helping to identify objects as well as their context (Goffaux et al., 2005; Rousselet, Joubert & Fabre-Thorpe, 2005; Torralba, 2009). Rivest and Cavanagh (1996) found that contours as defined by luminance, colour, motion and visual texture were equally weighted and integrated when establishing discrete objects. However they noted that "discontinuities in luminance created by shadows are not reliably linked to object contours, whereas continuities in other attributes (e.g. colour, motion and texture) are much more reliably linked to object contours.” Despite visual perception using a conjunction of multiple forms of information to define object boundaries, this is not possible with the most commonly researched SSDs.

The use of luminance alone in SSDs means that luminance-boundaries need to also be object-boundaries. This results in experimental stimuli for greyscale SSDs such as the ‘vOICe’ or ‘Brainport’ using single colour high contrast objects and backgrounds under consistent lighting conditions for object recognition and navigation (Auvray, Hanneton & O'Regan, 2007; Chebat, Schneider, Kupers & Ptito, 2011; Segond et al., 2005 – see fig. 4.1). The use of greyscale SSDs in more natural environments have large limitations, in which an objects’ signal varies depending on camera positioning, light positioning, light intensity, object characteristics (e.g. dark or bumpy) alongside potentially complex background information not relevant to the user (Brown, Macpherson & Ward, 2011; Capalbo & Glenney, 2009).

Figure 4.1. Left side shows object discrimination stimuli used in Auvray et al., 2007; Right side shows navigation stimuli used in Chebat et al., 2011. Both experiments use high contrast black and white stimuli with consistent lighting conditions, aspects that are less assured in complex natural environments.

In contrast to greyscale approaches which feature one boundary for edges (dark-light), colour features two additional boundaries, namely saturation (grey-colourful) and hue (red, yellow, green, blue). These three boundaries can be utilised in any combination to extract an object from its background. In computer-vision studies, a combination of luminance and hue information has been found to be reliable for computer navigation and object recognition even in shadowed environments (Crisman & Thorpe, 1993; Orwell, Remagnino & Jones, 2001). Computerised object segmentation under variable lighting can be aided through using hue as a more stable object marker, this can also be used to reduce the influence of shadows similar to colour constancy mechanisms (Salvador, Cavallaro & Ebrahimi, 2004). Similarly in human perception, Gur and Akri (1992) found that luminance and hue processing was enhanced by the integration between these in the same stimuli (see also Syrkin & Gur, 1997). As such, the conjunction between multiple forms of colour information may be beneficial over those same cues in isolation.

Previous studies on colour in sensory substitution have used a wide variety of tasks that incorporate colour information, from simplistic abstract representations of colour for simple shapes or flags on a computer (Ancuti, Ancuti & Bekaert, 2009; Burch, 2012). It could be argued that understanding abstract colours through sound is more a measure of how well participants can represent distinct sounds with new labels. Interestingly this point has been well illustrated previously where colour information has even been presented without a visual context or application, where participants were able to obtain an understanding of a specific 'colour space' through how perceptually distinct two given points are from one another in this space even when the participants did not know it was colour that was being represented (Kahol et al., 2006). By comparison, some tasks have involved live feeds of artificially produced highly saturated colours in a real environment, such as through matching differently coloured socks (Bologna et al., 2008), navigating along painted lines outside (Bologna et al., 2010) or recognising coloured doors inside (Meers & Ward, 2004). Finally some devices have been used on naturally occuring colours such as through identifying a variety of fruit indoors (Capalbo & Glenney, 2009), the use of natural colours has the advantage of both being a part of the evolutionary basis from which colour vision originally developed (Jacobs, 2009) and introduces meaningful complexity in terms of shades of colour that reflect real environments. Another additional factor to consider is that the colours of objects vary according to their environmental lighting conditions so a 'red apple' will be a variety of shades of red as you move it across a variety of lighting conditions. This element of colour comprehension by colour SSD users has to date not been addressed.

There have been previous attempts to compare the use of colour and greyscale in sensory substitution; however these have come with methodological problems. Ancuti, Ancuti and Bekaert (2009) compared the colour based ColEnViSon and greyscale vOICe SSDs on discriminating the patterns of flags using four participants. Both SSDs performed similarly for questions of orientation and complexity; however the ColEnViSon had a marked improvement for identifying the number of colours present and recognition of specific flags. This is to be expected since similarly luminant colours would be indistinguishable on the vOICe, capping its performance. Besides the amount of information, the style of presentation to the user also differs, for instance, pitch denotes vertical location in the vOICe and luminance in the ColEnViSon. This makes the influence of different translations difficult to gauge. Similarly, Capalbo and Glenney (2009) compared their colour SSD, the Kromophone, which converts a single point of colour into sound with the vOICe for light localisation and fruit recognition tasks. Their preliminary findings using three to six participants found that the Kromophone was superior for localisation and recognition as well as more resistant to changes in environmental illumination. However the Kromophone and vOICe are also difficult to compare because they convert different spatial dimensions and use different translations for luminance information making it difficult to conclude whether differences in information content or translation underpin these differences. The present experiment seeks to address these concerns by using the same SSD to convert either luminance or colour information, as well as present this information using a variety of representations. From this it becomes possible to disentangle the influence of information content from information presentation to the user. The next section examines the variety of colour-spaces relevant to this experiment as well as potential ways this can be turned into sound.

4.2.2. Designing colour for SSDs

Across previous colour SSD designs, a wide variety of approaches have been taken (Hamilton-Fletcher & Ward, 2013). One of the ways in which these devices vary is in their representation of spatial dimensions, with vertical position indicated through pitch or channel, horizontal position through timing, panning and inter-aural differences (Ancuti et al., 2009; Bologna et al., 2007; Levy-Tzedek et al., 2012a, 2012b). Sound-space representations limit the number of auditory dimensions free to represent colour. Some devices avoid this by having the user select a single point in space for sonification, such as the Soundview SSD (Doel, 2003). The Soundview uses the spatial position of a stylus on a tablet to sonify the colour in that location. Over time the user is able to map out the distribution of colours over the entire image and concentrate on areas of difficulty. Using this approach also means that the auditory representations remain neutral with respect to spatial mappings to concentrate entirely on representing colour in sound. If colour is represented in SSDs, first a colour space needs to be chosen before figuring out a specific mapping of these visual dimensions into sound. These questions are explored in turn in the following sections.

4.2.2.1. Varieties of colour information

While greyscale devices such as the 'vOICe' and 'Brainport' both use a single dimension of luminance, colour SSDs can represent colour in a variety of ways. Devices such as the Soundview, See ColOr and original Kromophone use variations of HSL colour space that supplement luminance information with dimensions representing saturation and hue (Bologna et al., 2007; Capalbo & Glenney, 2009; Doel, 2003). Coding these dimensions as independent from one another requires the use of polar co-ordinates (see fig. 4.2). As such, while dimensions for luminance and saturation vary in a linear fashion (low-high), hue becomes a circular dimension typically given in degrees, cycling between the hues. Unlike luminance and saturation, values do not necessarily correspond to similarity, as hues far apart on this dimension may be perceptually similar. This considered, the use of independent dimensions may still allow for a more intuitive understanding of colour, potentially aiding in colour identification (Berk, Kaufman & Brownston, 1982; Schwarz, Cowan & Beatty, 1987). An alternative approach to independent dimensions is that of using related dimensions and Cartesian co-ordinates (see fig. 4.2), where variations on one dimension vary a combination of hue, saturation and luminance. One example of this is RGB space which is typically seen in computer models of colour. This consists of three dimensions relating to the saturation of red, green and blue. Varying their combination changes the final hue (e.g. yellow is a combination of red and green), saturation (saturated colours have unequal R, G and B values, while greyscale is equal) and luminance (higher R, G and B values are more luminant). In sensory substitution the tactile-colour glove (Cappelletti, Ferri and Nicoletti, 1998) translates the R, G and B dimensions to the ring, middle and index finger respectively through increased vibrotactile stimulation. Despite HSL and RGB space using different approaches to colour, each allows the ability to identify a variety of colours and contours as defined by changes in luminance, saturation and hue.

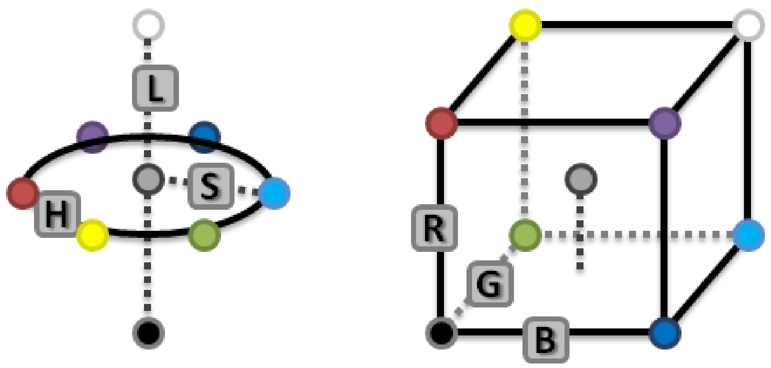


Figure 4.2. HSL (left) and RGB (right) colour spaces. HSL colour space uses three independent dimensions relating to the hue, saturation and luminance of the final colour. RGB colour space uses three dimensions indicating the level of red, green and blue saturation. Mixes of these dimensions change the final colours' hue, saturation and luminance.

4.2.2.2. Approaches to representation in sound and touch

After the type of colour space is chosen, the colour's representation in sound and touch is the next factor to consider. There are several aspects that constrain these choices or suggest optimal mappings. The first limitation is that the total amount of information that can be discriminated at a given time is lower for hearing and touch than it is for vision (Jacobson, 1951a, 1951b; Kokjer, 1987). This upper limit for each sense creates a perceptual bottleneck, so visual information would need to be reduced (e.g. reducing spatial dimensions) before translation. Another limit is the range of dimensions that can be used for each modality. For example hearing could utilise variations of pitch, loudness and interaural differences, while touch could use variations of frequency, intensity and location to represent visual dimensions. The specific mappings between the senses can draw upon widespread intuitive pairings known as cross-modal correspondences, where congruent pairings (e.g. high pitch and high spatial location) are processed faster than incongruent pairings (Spence, 2011). It is not known if all correspondences work equally well, for example, since luminance corresponds to auditory pitch, loudness and vibrotactile frequency (Lewkowicz & Turkewitz, 1980; Martino & Marks, 2000; Ward et al., 2006) is there an optimal representation? Saturation has been mapped to harmonics and loudness (Giannakis, 2001; Ward, Huckstep & Tsakanikos, 2006). For hue, blue and yellow have been matched to low and higher pitches respectively (Simpson, Quinn & Ausubel, 1956), however this might be confounded by variations of luminance (Spence, 2011). As such there are a variety of correspondences which can be utilised to construct colour spaces.

4.2.3. The Creole – turning colour into sound and touch

In order to compare different approaches to colour and sound representation we utilised 'the Creole,' a touchpad sensory substitution device that turns visual images into patterns of sound and touch (see fig. 4.3).

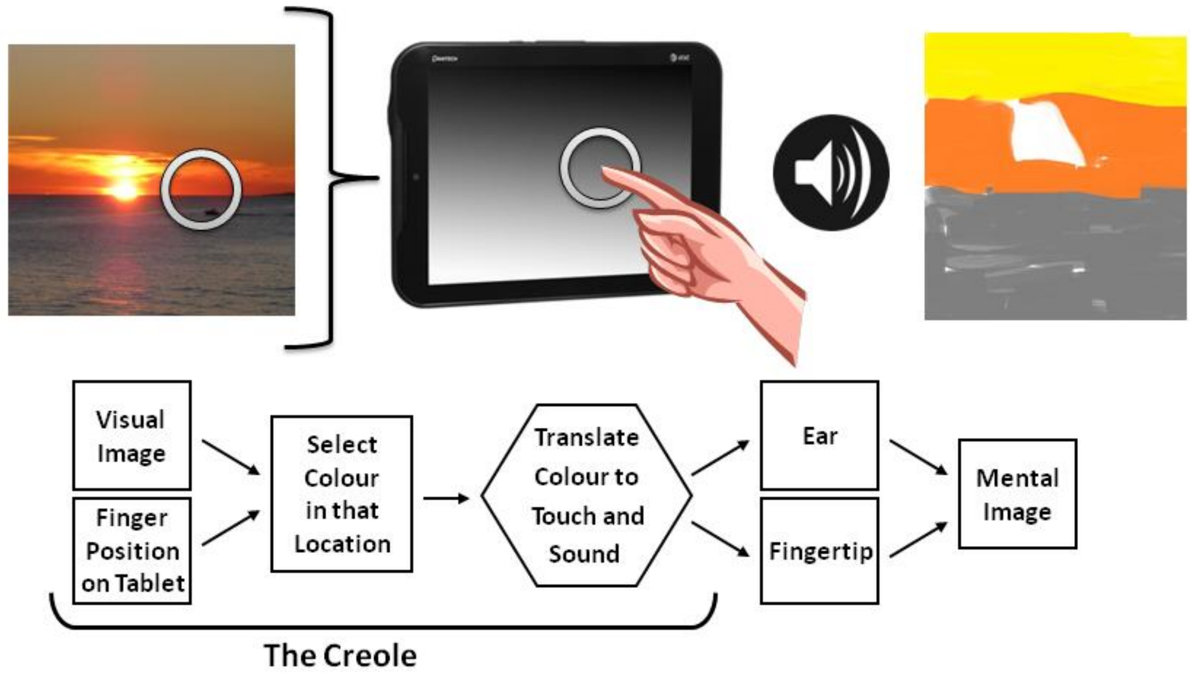


Figure 4.3. The Creole. Visual images are stored on the device but not shown to users. Touching the device relays co-ordinates of the finger to its equivalent position on the image and selects the pixel at that location. The pixel's colour values are then translated into patterns of touch and sound for the user. Through explorations over time, the user can build up a mental image of the distribution of colours in the image.

The Creole uses one of four mappings, two of which only use luminance information and two that code colour. The two greyscale mappings are 'luminance-pitch' and 'luminance-loudness,' while the two colour mappings are 'hue-pitch' and 'RGB-loudness.' These are described in more detail below as well as their implications for how luminance, saturation and hue boundaries are given to the user.

4.2.3.1. Greyscale mappings

Luminance information was used from HSL colour space, which represents a value ranging from 0 (black) to 100 (white). For the 'luminance-pitch' mapping, the luminance dimension is logarithmically mapped to auditory frequency ranging between 500Hz (black) and 5000Hz (white). This is because pitch is perceived logarithmically relative to frequency. The frequency range was chosen due to its effective use in SSDs such as the vOICe (Meijer, 1992) and covering the frequencies most sensitive to variations of loudness (Gelfand, 2011). For the 'luminance-loudness' mapping, the luminance value was logarithmically scaled to the amplitude of three outputted pitches (523, 659 & 784Hz) which are all heard simultaneously at equal volume. The choice of these pitches is based on the 'RGB-loudness' mapping detailed below so that similar sounds can be compared with different amounts of colour information.

4.2.3.2. Colour mappings

Colour information is represented in one of two ways in the Creole. The 'hue-pitch' mapping uses HSL colour space, logarithmically mapping the hue dimension (0-360°) to auditory frequency ranging between 500 and 5000Hz, starting at blue and progressing through cyan, green, yellow, red and purple. Saturation and luminance are mapped to auditory amplitude and vibrotactile amplitude respectively. This mapping uses previously observed hue-pitch, saturation-loudness and vibration-luminance correspondences (Giannakis, 2001; Martino & Marks, 2000; Simpson et al., 1956). So a dark saturated blue would produce a loud low-pitched tone with low vibration intensity. The second colour coding, 'RGB-loudness' associates the blue, green and red dimensions with the pitches of 523, 659 & 784Hz respectively, while saturation along these is mapped to their amplitude. These pitches in combination make a major C chord, giving a harmonic quality to saturated values while placing blue with lower pitches and yellow (red + green) with higher pitches. These mappings mirror saturation-loudness, saturation-harmonics and hue-pitch correspondences (Giannakis, 2001; Simpson et al., 1956; Ward et al., 2006). An argument could be made that since the 'hue-pitch' coding uses a luminance-vibrotactile mapping, any affect on performance could be the result of the tactile feedback rather than the auditory feedback. To counter this criticism, all other codings ('luminance-pitch,' luminance-loudness' and 'RGB-loudness') were given this luminance-vibrotactile mapping so there was no variation in tactile stimulation across all groups. This results in participants across all groups, access to luminance information through the tactile feedback.

4.2.3.3. Detecting boundaries

An important consideration for end users is how changes in hue, saturation or luminance in an image are perceived as a result of the different mappings. Greyscale encodings can only perceive changes in luminance through either variation of pitch or loudness, as well as vibrotactile stimulation. Changes in hue and saturation can only be perceived by the colour encodings 'hue-pitch' and 'RGB-loudness' through variations of pitch and loudness. Luminance is reflected in the overall loudness of the three pitches in 'RGB-loudness,' while both approaches represent luminance in vibrotactile intensity as well. Table 4.1 provides an illustration of how each boundary is experienced across the mappings.

Table 4.1. Change in stimulation experienced as users cross a hue, saturation or luminance boundary.

Boundary	Lum-to-Pitch	Lum-to-Loudness	Hue-pitch	RGB-Loudness
Hue	n/a	n/a	Change in pitch	Change in loudest pitch
Saturation	n/a	n/a	Change in loudness	Change in individual pitch loudness
Luminance	Change in pitch, vibration	Change in loudness, vibration	Change in vibration	Change in overall loudness, vibration

4.2.4. Hypotheses

The present experiment seeks to assess the effect of information type (luminance, colour) as well as different representations of this information (e.g. luminance-pitch, luminance-loudness) on discrimination in tasks involving colour. In order to test these factors, a fruit discrimination task across differing environmental luminances was used as this features a natural use of colour and is potentially solvable with greyscale approaches. Finally in order to assess colour knowledge, a colour replication task was employed, where participants would attempt to visually recreate colours that have been heard through the Creole.

For the object discrimination tasks, we predict that encodings that sonify colour will outperform greyscale approaches as seen in visual tasks (Crisman & Thorpe, 1993; Orwell, Remagnino & Jones, 2001; Gur & Akri, 1992; Rivest & Cavanagh, 1995; Torrabla, 2009). We also predict that there will be fewer errors due to mistaken objects (e.g. mistaking an orange for an apple) for colour approaches, while errors of luminance (i.e. mistaking bright and dark environments) will remain stable across encodings. Since both greyscale mappings utilise correspondences we expect both to perform equally.

For the colour replication tasks, we predict that representations using independent dimensions such as ‘hue-pitch’ will provide less error than ‘RGB-loudness’ since it is easier to decode independent dimensions (Berk et al., 1982; Schwarz et al., 1987). For greyscale approaches, we expect errors of luminance to be equal across encodings, as both mappings utilise correspondences.

4.3. Method

4.3.1. Participants

A total of 53 participants (Mean age = 22.28, SD = 6.18, 9 male, 4 left-handed) were recruited from the University of Sussex using the online recruitment service Sona. The education of the participants included four studying at post-graduate level with the remaining at undergraduate level. All participants had normal or corrected-to-normal vision and hearing. None of the participants reported colour vision deficits or the presence of sound-colour or touch-colour synaesthesia. Participants were able to choose their reimbursement which was either course credits for their psychology degree or payment of ten pounds. All participants gave informed consent and were randomly assigned to their Creole-coding group.

4.3.2. Design

For the object discrimination task a between groups design was implemented consisting of four groups relating to Creole-coding with the dependent variable of percentage fruit/vegetables correctly identified during a four alternate forced choice task. The four independent groups consisted of the 'pitch-luminance,' 'loudness-luminance,' 'hue-pitch' and 'RGB-loudness' encodings. The four choices given in the 4AFC task to participants consisted of the target object in a high and low environmental luminance, as well as a distracter object in high and low luminance. Only the target object in the same environmental luminance condition that was turned into auditory/tactile stimulation on the Creole was considered a correct answer. Three further analyses were done comparing the types of errors made (luminance-error, object-error or both) across groups. For the colour replication task, two separate between groups analyses were carried out either comparing the colour ('hue-pitch,' 'RGB-loudness') or greyscale ('pitch-luminance,' 'loudness-luminance') approaches, since these approaches explore different amounts of colour-space and potential error.

4.3.3. Materials

4.3.3.1. The Creole

The 'Creole' is a sensory substitution device which allows users to explore a hidden visual image on a touchpad with their finger, turning the colours underneath their fingertip into auditory and tactile feedback. Over time the user can learn the location of colours in a serial fashion to build up their understanding of this hidden visual image. Visual images are loaded into the Creole program allowing the participant to select individual pixel colours through the centre-point of their finger, while lifting the finger off of the screen stops the auditory and tactile output. Participants can stroke the screen to hear the progression of individual pixels over time, or dab the screen to quickly hear individual pixels immediately. When images are loaded onto the Creole, they take up the whole screen, the left third only ever contained reference information during training, and the right third contained target information during all tasks while the middle third was black separating the two. The device is named after creole languages where a new language is formed from two parent languages; likewise the 'Creole' SSD is the result of using both visual-auditory and visual-tactile mappings together.

The touchpad used was a pantech 4G LTE running the Creole client program on android OS while connected via USB to a Lenovo T520 laptop that runs the application that informs the Creole client program. Each picture and colour-coding method is associated with a separate application running, therefore between each image, the completed image application was shut down and a new image application was started. The starting of a new image application results in a starting tone that allows the user to know there is a new image on the device for them to explore (since no visual feedback is provided on the device itself). The online processing of audio and tactile output from finger explorations over the image was done using an Intel HD graphics 3000 GT2+ graphics card. Audio stimulation was generated from the laptop program using a Conexant 20672 SmartAudio HD soundcard outputted through commercial earbuds. Tactile stimulation was generated from the tablet itself with increasing luminance of individual pixels resulting in increased vibrotactile feedback. The program for the tablet and laptop was coded in C# using Microsoft Visual Studio 2012.

The equation for the 'luminance-pitch' and 'hue-pitch' mappings used the following formula to logarithmically map the proportion of luminance or hue ('colour' ranging 0 to 1) to an auditory frequency between 500 and 5000Hz:

$$X = colour / (\log_2 (5000/500))^{-1}$$

$$Frequency = 500 * 2^X$$

The proportion of saturation (0 to 1) for the 'hue-pitch' coding was silent for values under 0.3, so white, grey and black were completely silent, while values over 0.3 were logarithmically transformed into auditory loudness with the following formula:

$$Frequency\ amplitude = ((saturation * saturation) / (1 * 1))$$

The proportion amplitude for each of the three frequencies in the luminance-loudness and RGB-loudness encodings was worked out as follows, which each pixel colour value (R, G & B) ranging between 0 and 255:

*Frequency amplitude = ((pixel * pixel) / (255 * 255))*

The proportion of amplitude for vibrotactile stimulation was directly mapped to the proportion of luminance in HSL space for all encodings.

4.3.3.2. Training materials

The first part of the experiment required that the participants could effectively use the Creole device to make judgements on hidden visual images by using the auditory and tactile feedback from the Creole. These hidden images used for training on the Creole consisted of 1080 by 720 pixel images (see fig. 4.4), which contained a palette of colours on the left half, and a target image for training on the right half. Participants could never view these images visually and were restricted to only exploring them through using the auditory and tactile feedback of the Creole. In the images used for training, the leftmost 220 by 720 pixel block contained a wide variety of hues arranged vertically in the image (starting with blue at the bottom, cycling through the hues to purple at the top), while different levels of saturation are presented horizontally (most saturated on the left, less saturated on the right). The next block in the image was also 220 by 720 pixels, and illustrates variations of greyscale from black on the bottom, through grey to white at the top. A 200 by 720 pixel black border separated these reference slides and the 440 by 720 target image which was located on the right. While participants could not see any visual images on the Creole itself, it was important that participants knew the distribution of colours on the Creole's reference side so that they could explore how variations in hue, saturation and luminance change the auditory and tactile feedback. To give participants this information, a picture of the reference side's colours was given to participants.

During the initial colour training on the Creole, the target images included pure blocks of colour (black, grey, white, red, green, blue, yellow, purple, cyan or orange) presented one at a time for participants to explore, progressing to the subsequent colour after participants were confident with how the colour sounded on the Creole. The groups that had access to all dimensions of colour on the Creole ('hue-pitch,' 'RGB-loudness') went through all of the coloured blocks, while those that only had access to greyscale ('luminance-pitch,' 'luminance-loudness') only explored the black, grey and white blocks as all other colours would provide the same stimulation. After the colour training, shape training blocks were also presented sequentially to participants, consisting of red target shapes (a circle, diamond or square) on a brown background. The shape training stimuli were chosen to prepare users for different shaped objects with a brown background, mimicking the brown table used for the fruit/vegetable discrimination stimuli used later.

During the colour and shape training, visual examples of the targets they were listening to (either all the blocks of colour or all the shapes) were presented simultaneously on a separate computer's monitor. This was so that participants had a visual example of what could be being represented on the Creole before making their choice on which colour or object they believed was being represented on the Creole.

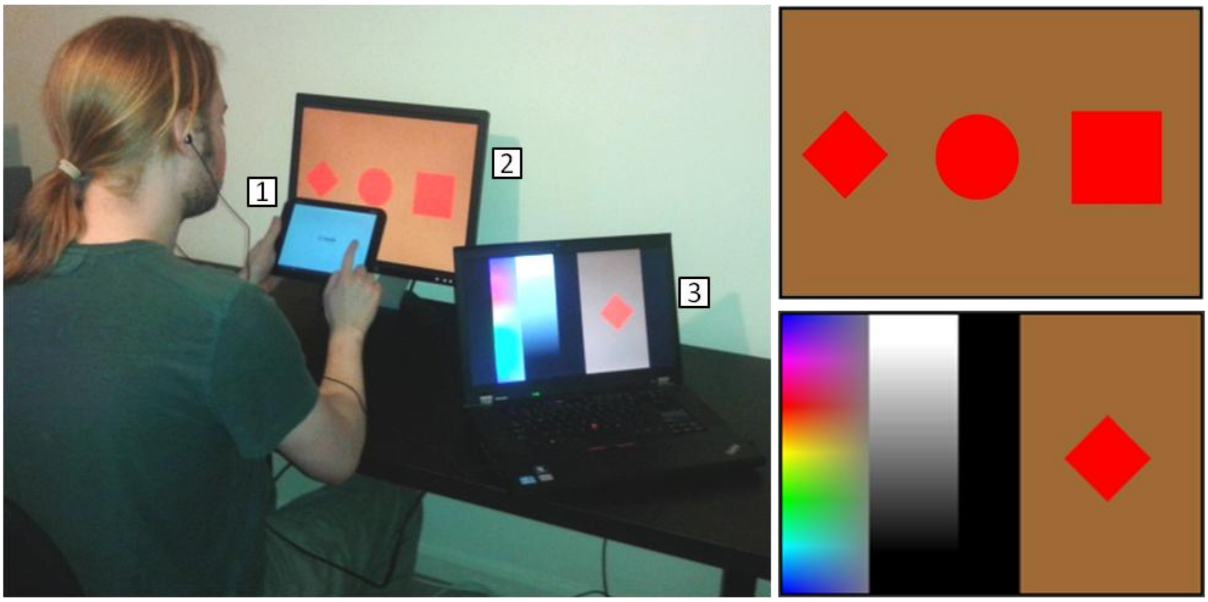


Figure 4.4 – Use of training images for the Creole. Left image shows experimental set up during all tasks. The Creole is held by the participant and gives no visual feedback to the user (1), the location touched on the Creole corresponds to the same spatial location on the hidden image (see bottom right image) that is run on a connected laptop (3) that is never seen by the participant. A monitor displays the available options for the participant to choose between (see top right image). Top right image shows the shape options visually presented to the user in order to select which they believed to be in the 'target region' of the Creole. Bottom right image shows the hidden shape training image on the Creole, the left side of this allows an exploration of variations in hue, saturation and luminance for the user, whereas the right side contains the target object.

4.3.3.2.1. Object discrimination task

For the object discrimination task, participants are tasked with figuring out which specific fruit or vegetable is being represented through auditory / tactile stimulation on the Creole from a choice of four that are being visually presented to them. Each target fruit or vegetable listened to on the Creole was contained within the rightmost 440 by 720 pixel section of the Creole. These hidden images featured a photo of a single fruit/vegetable on a light brown wood grain background.

Different to the training session, during this task there was no reference set of colours to listen to on the left. Instead, the remaining pixels for the hidden image on the Creole were black.

The visual stimulation for the participants indicated the four objects participants could choose their answer for in a four alternate forced choice task. The four choices always consisted of two fruit/vegetables, consisting of the randomly chosen target image that the participant would be listening to on the Creole and a randomly chosen different distracter fruit/vegetable. The varieties of fruit/vegetables available for sonification reflected the colours that were used in the colour training section, with fruit/vegetables primarily centred around green (lime, pepper), orange (orange, carrots), red (apple, pepper), yellow (lemon, pepper) and purple (onion, plum). Each image was either represented in their natural luminance (referred to as 'bright') or the luminance was reduced by 50% using gimp image editing software (referred to as 'dark') creating 20 possible targets and distracters. The visual presentation to the participants consisted of the target object (e.g. bright apple) alongside distracters consisting of a change in luminance (e.g. dark apple), change in object (e.g. bright orange) or both (e.g. dark orange) in a 2 by 2 matrix (see fig. 4.5). The bright variants were in the top two locations and dark in the bottom two locations. The choice of targets and distracters were randomised in Microsoft Excel. Each participant went through 20 fruit/vegetable discrimination trials in a sequential manner. In each trial participants take a minute to listen to the randomly selected target object and make their selection from the four available options.

The visual presentation was displayed using Microsoft PowerPoint which displayed a black screen until any key is pressed and then displays the 2x2 matrix of target and distracter objects. This image is displayed for 1 minute until it automatically progresses to the next black blank screen. This blank period allows the participant to give their answer and allow the next target image to be loaded onto the Creole by the experimenter. The PowerPoint presentation was presented on a 19 inch flat screen monitor running at 1024 by 1080 resolution with 60Hz refresh rate.

Figure 4.5 – Object discrimination stimuli. Left side shows the hidden image represented on the Creole, the target image takes up the right side of the screen with no reference information given on the left side. The right image shows the 2x2 matrix of target and distracter images that is visually presented to the participant for selection.

4.3.3.2.2. Colour replication task

The final task examined how well participants understood the colour sonifications they were listening to on the Creole. This was done by tasking participants with recreating the colours they believed they were listening to. For the colour replication task a 2 by 5 matrix of randomly selected colours was represented in the target area on the right side of the Creole (see fig. 4.6). Participants were never able to see these colours visually and could only listen to them. For the Creole groups that coded saturation and hue information (the 'hue-pitch,' and 'RGB-loudness' groups), each of the 10 colours in the matrix was selected through using three randomly selected values between 0 and 255 for each of the R, G and B dimensions of RGB-space in Microsoft Excel. These values were combined in RGB-space to create the randomly selected final colour for each section of the matrix. For the greyscale Creole codings, a single random value between 0 and 255 was applied to R, G and B dimensions together as equal RGB values only vary the luminance of a colour between black (0, 0, 0) and white (255, 255, 255) while saturation remains at zero. All 10 colours were presented simultaneously, so users were able to explore the colours represented on the Creole in any order for comparisons. The responses from participants were given into a similarly shaped 2 by 5 matrix of blank white boxes, which the participants were able to edit each box's colour using the advanced colour selector option in Microsoft Paint tools. The colour palette available to participants was a selection based on HSL space giving a selection of hues and saturations with a luminance slider. Manipulation of this colour provided both the HSL and RGB co-ordinate information as well, so there was no unique advantage given to the 'hue-pitch' or 'RGB-loudness' groups. For the greyscale Creole coding groups, participants were reminded that their coding did not involve saturation or hue information and therefore choosing a colour other than greyscale would not make sense. Participants were tasked with trying to recreate the colours they had heard on the Creole as real colours that would be used to fill in the spatially corresponding box (so the top right section on the Creole would correspond to the top right section in the Microsoft Paint image).

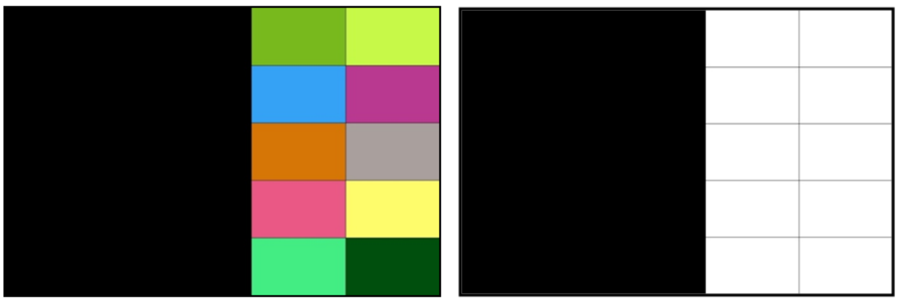


Figure 4.6 – Colour replication stimuli. Left side shows a hidden image represented on the Creole, the image consists of a 2 by 5 matrix of randomly selected colours which take up the right side of the screen with no reference information given on the left side. The right image shows a blank 2x5 matrix presented on a separate computer for the user to fill in the colours they believed they were listening to from the Creole.

4.3.3.2.3. Image illustration

A final optional task was created for motivated participants to attempt if time permitted. This task involved a complex natural image (e.g. street scene, sunset or painting) being represented on the Creole with participants attempting to recreate this using the Microsoft Paint tools. Participants were told this optional task may take up to 30 minutes prior to committing.

4.3.4. Procedure

Participants were recruited using the Sona online research advertisement system at the University of Sussex, the experiment was described as ‘doing a series of visual tasks using sound.’ During the testing session, participants were given an information sheet before filling out a consent form and demographics sheet. Then participants sat down at a computer displaying a series of colours and were handed a sheet giving a written description of the Creole device and the encoding method used for their group. The experimenter also verbally described the device and translation method and participants were encouraged to rest the device in their non-dominant hand in order to feel the tactile vibrations at all times. Participants were informed that only one pixel is turned into sound at a time and that dabbing the screen with their finger kept the latency between finger press and auditory feedback low, while longer contact periods did unfortunately increase the latency. At this point participants were informed of the starting sound that indicates a new image has been loaded onto the device, before the first training image was loaded. The first image features a single shade of colour that could be sonified by the user through touching the right side of the Creole. This image also featured a reference set of colours on the left allowing an exploration of hue, saturation and luminance. The reference side on the Creole is identical to the reference side shown in the written description of their Creole (see fig. 4.4). Participants were encouraged to explore the reference side to listen to all the colours and match up the ‘target colour’ on the right side, with reference colours on the left side in order to name the target colour. Participants were then taken through the remaining colours used to select the fruit/vegetables in the main task that the device could discriminate. After the colour training participants were able to try three shape discrimination tasks through exploration, choosing what shape they believed they were listening to from three choices on the computer screen. Participants were given feedback during all the training tasks and overall the training took approximately 15 minutes.

The object discrimination task was then presented to the participants. They were told they would be presented with four images on their computer screen consisting of two different types of fruit/vegetable, both of which are either in high or low luminance lighting conditions. Participants were told that one of the four images would be presented on the Creole device to explore and listen to, and that they would have one minute to determine which of the four images they were listening to. When the image was loaded onto the Creole, participants heard the starting sound, whereby they pressed the spacebar to present the four images and started exploring the Creole image. After a minute the screen went blank and participants were asked which image they believed they were listening to before the next image was loaded onto the Creole. After this minute, participants were not allowed to further explore the hidden image on the Creole but could deliberate for another 30 seconds before being pressed for an answer. If participants responded faster than a minute, participants were still able to explore and change their answer when the minute was up. This process was repeated 20 times for each participant. No feedback was given to participants during this task.

After the object discrimination task, participants were presented with a blank 2 by 5 matrix and were told a similar matrix would be sonified on the Creole device but with a randomised colour within each box and that it was their task to listen to the colours and recreate them in the blank matrix on the screen. Participants were walked through using the advanced colour selection options and fill options in Microsoft Paint. The colour selection options included a colour swatch of HSL colour space with access to RGB co-ordinates. Participants were given an unlimited amount of time to make their selections and could go back and change their colours at any point. Afterwards participants were debriefed as to the nature of the experiment, reimbursed and thanked for their time. If there was a free subsequent time slot, participants who had been trained on a colour coding for the Creole were offered an optional image recreation task. Participants were told this may take up to 30 minutes, that there was no pressure to commit, no further reimbursement and that they could stop at any time.

4.4. Results

After group scores were collected, participants that scored beyond two standard deviations of their group were excluded as outliers. This criterion was chosen due to the relatively small group sizes so that no individual score had exaggerated influence on group means.

4.4.1. Object discrimination task

The amount of correctly identified objects from twenty trials was collected for each group. With four options and one correct answer, the chance performance level was five correctly identified objects. The greyscale mappings of luminance-pitch and luminance-loudness had mean group scores of 9.09 (SD = 1.58) and 11.61 (SD = 2.75) respectively, while the colour mappings for hue-pitch and RGB-loudness space had mean scores of 7.58 (SD = 1.68) and 11.38 (SD = 2.29) respectively. A between groups ANOVA was performed and this revealed a highly significant main effect of group, F (3, 45) = 9.86, p < .001. This pattern was further analysed with bonferroni-corrected post-hoc tests, which indicated significantly higher scores for the luminance-loudness mapping over both luminance-pitch (p = .039) and hue-pitch mappings (p < .001). Likewise the RGB-loudness space mapping showed significantly higher object discrimination scores over hue-pitch space mapping (p < .001). The RGB-loudness mapping also trended towards significance in comparison to the luminance-pitch mapping, however this was non-significant (p = .077). The post hoc tests indicated that mappings that utilised loudness for luminance (present in luminance-loudness and RGB-loudness mappings) performed better than modes that utilised a broad range of pitches for luminance or hue (see fig. 4.7). Colour information did not have a uniformly positive effect on object discrimination, with no significant difference between luminance-loudness and RGB-loudness or luminance-pitch and hue-pitch.

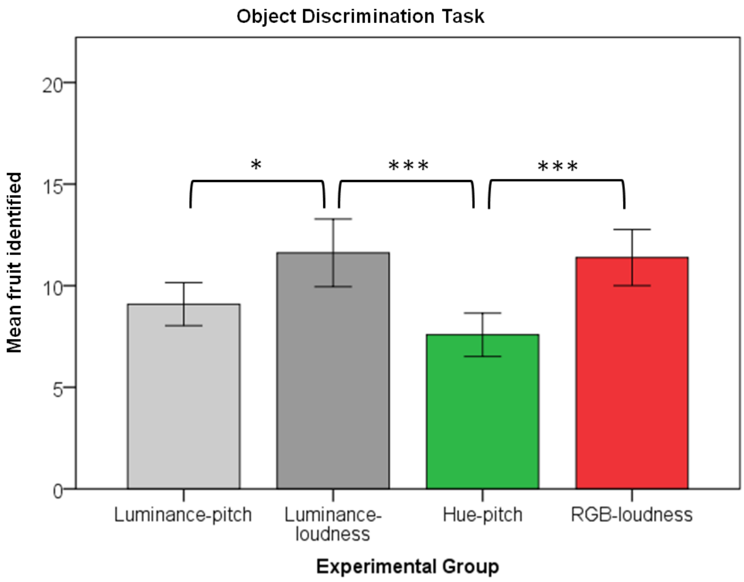


Figure 4.7 – Mean correctly identified targets in the object discrimination task. Left of centre are greyscale encodings, right of centre are colour encodings. Luminance- loudness significantly outperformed luminance-pitch and hue-pitch mappings, while RGB-loudness significantly outperformed the hue-pitch mapping. Error bars indicate 95% confidence intervals. Key: * = p <.05, *** = p <.001.

A subsequent analysis examining the types of errors was conducted. For example if the participant had a 'bright apple' represented on the Creole mistakes were categorized into object errors (i.e. saying "bright *orange*"), luminance errors (i.e. saying "*dark* apple") or a combination of both (i.e. saying "*dark orange*"). An independent ANOVA was carried out, using all three types of error as separate dependent variables. This analysis revealed that there was no significant main effect of group for object errors, F (3, 45) = 0.77, p = .515 or combination errors, F (3, 45) = 1.51, p = .224. The examination of luminance errors found that there was a violation of homogeneity of variances across groups therefore Brown-Forsythe's F-Tests are reported. It was found there was a main effect of luminance errors across the Creole encoding groups, F (3, 37.91) = 7.69, p <.001, further bonferroni-corrected post hoc tests found that the RGB-loudness mapping committed significantly fewer luminance errors than the hue-pitch mapping (p = .026), similarly the luminance-loudness mapping group committed fewer luminance errors than both luminance-pitch (p = .010) and hue-pitch mappings (p = .001). Therefore it appears as if the improved object discrimination abilities are underpinned through luminance being encoded in sound as loudness, a trait present in both RGB-loudness and luminance-loudness mappings.

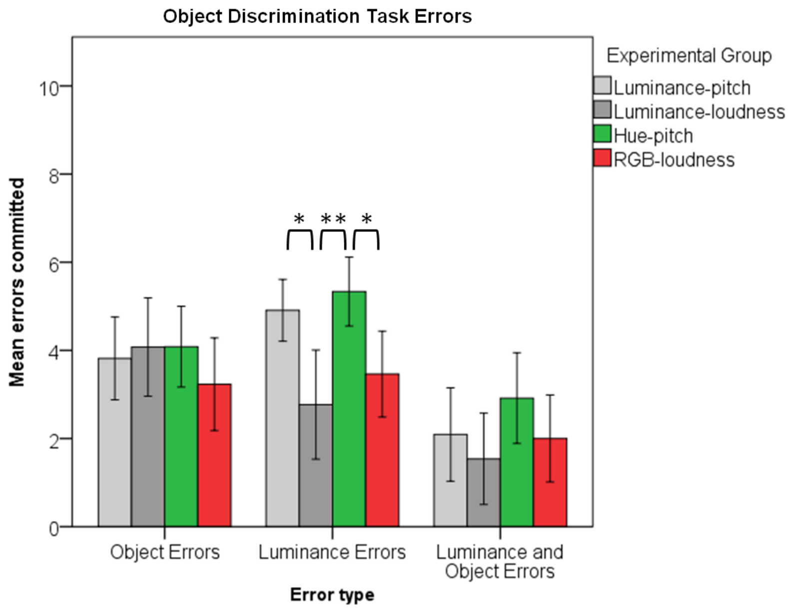


Figure 4.8 – Average number of object errors, luminance errors and combination errors. Left of error groupings are greyscale encodings, right of error groupings are colour encodings. Luminance to loudness had significantly less errors than luminance to pitch and hue-pitch mappings, while RGB-loudness had significantly less errors than the hue-pitch mapping. Error bars indicate 95% confidence intervals. Key: * = p < .05, ** = p < .01.

4.4.2. Colour replication task

Participants were presented with ten hidden colours on the Creole and asked to recreate them using the colour selection tools in Microsoft Paint. The level of participant error is determined by the distance in colour space between the Creole's colour point and the user's recreated colour point, with larger errors implying a worse understanding of the colour. Distances were calculated using CIE LUV colour space, which reflects perceptual distances to the human observer consisting of luminance (L), red-green (U) and blue-yellow (V) colour dimensions. The Euclidian distance between the Creole colour (CIE L1, u1 & v1) and participants' recreated colours (CIE L2, u2 & v2) were worked out using this formula:

$$\sqrt{((L1- L2)^2 + (u1- u2)^2 + (v1-v2)^2)}$$

A between groups analysis comparing the RGB-loudness mapping (Mean Error = 71.14, SD = 17.28) and hue-pitch mapping (Mean Error = 58.37, SD = 9.12) was conducted, revealing a significant difference in error scores for participants t (22) = -2.20, *p* = .038. This indicates that participants listening to the hue-pitch mapping were significantly more accurate in their final colour selections than those using the RGB-loudness mapping.

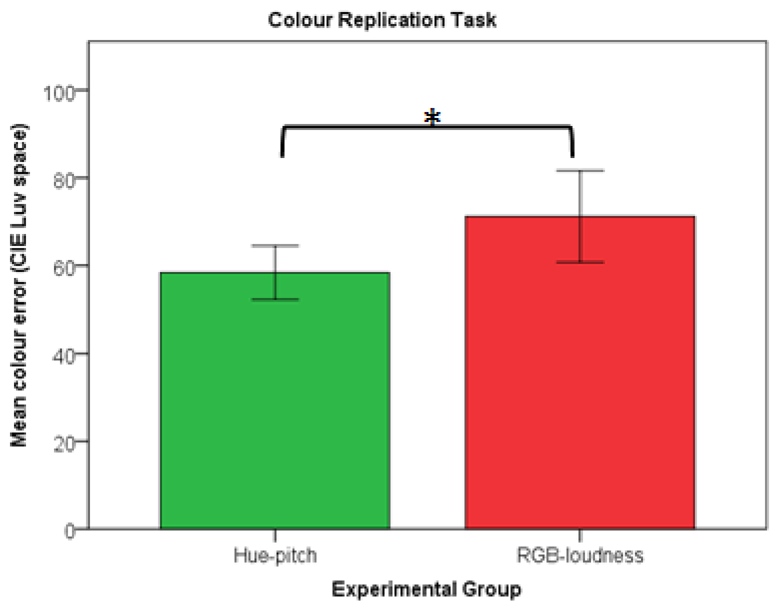


Figure 4.9 – Mean error score for the colour replication task in CIE LUV space. The hue-pitch mapping had participants making significantly less colour error than the RGB-loudness mapping. Error bars indicate 95% confidence intervals. Key: * = p < .05.

Differences in colour error for replications of greyscale colours for the luminance-pitch and luminance-loudness groups were also analysed. These were analysed only using the luminance dimension of CIE LUV space as any selections on hue or saturation would have been arbitrary choice for participants. The colour error was worked out using the following formula:

$$\sqrt{\left((L1 - L2)^2 \right)}$$

Results indicated that there was no significant group difference in the participant's recreations of greyscale colour, $t\ (23) = .55$, $p = .586$, with luminance-pitch having a mean of 20.9 (SD = 7.12) and luminance-loudness having a mean error of 19.45 (SD = 6.15).

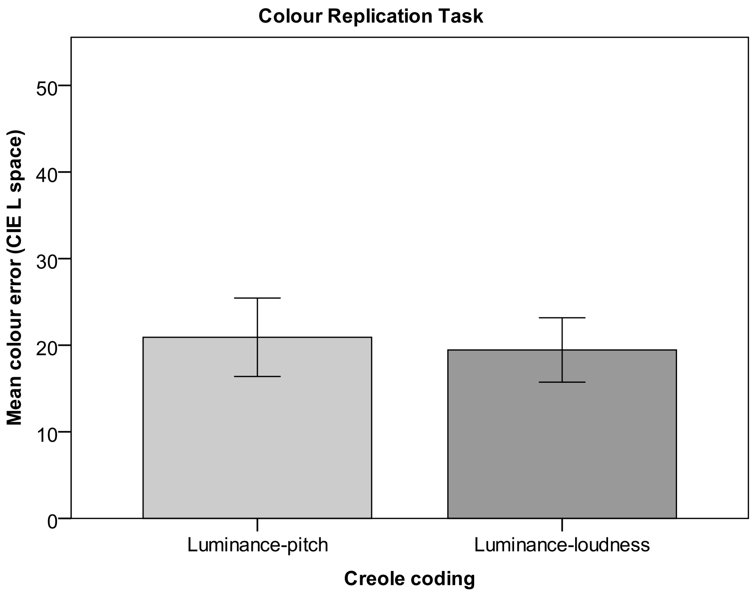


Figure 4.10 – Mean error score for the colour replication task in CIE L space. The luminance to pitch and luminance to loudness mappings performed at equal levels. Error bars indicate 95% confidence intervals.

4.4.3. Image recreation

After the behavioural tasks, motivated participants who had been trained in one of the colour encodings were offered an opportunity to attempt to recreate a complex natural image represented on the device. Participants were given 30 minutes to complete their illustration in Microsoft Paint. Some examples of the stimuli and participant illustrations can be seen in figure 4.11.

Figure 4.11 - Recreations of natural images hidden on the Creole by participants. Left – Illustration by participant using hue-pitch mapping recreating a street scene. Middle – Illustration by participant using RGB-loudness mapping recreating a sunset. Right – Illustration by participant using hue-pitch mapping recreating a Rothko painting.

4.5. Discussion

The present experiment examined the influence of colour information and its representation in sensory substitution during object discrimination and colour replication tasks. It was found that for object discrimination, colour information did not in of itself infer an advantage; instead the representation of information appeared to have the strongest influence on performance. Mappings that represented luminance information in loudness performed best, and this occurred primarily due to fewer errors made in identifying environmental luminance. Interestingly, not all correspondences performed equally well, with luminance-loudness outperforming luminance-pitch mappings. The colour replication task revealed that both luminance representations had an equivalent understanding of greyscale colours, meaning that this was not translated equally well into the more naturalistic object discrimination task. Finally the colour encodings that had the most accurate understandings of colour did not perform best at the object discrimination task indicating that more accurate colour knowledge does not necessarily improve performance in naturalistic tasks involving colour.

4.5.1. Representing colour effectively

For the object discrimination task, participants had to choose which of four fruit/vegetables was represented on the Creole with three distracter objects that varied in luminance, shape and colour. An analysis of errors indicated that the amount of errors due to misidentifying the fruit/vegetables (irrespective of luminance variations) was equal across all encodings. This indicates that the presence of colour information did not aid users in using colour cues to effectively eliminate differently coloured fruit/vegetables from their selections, so why might this be? Firstly, the hue-pitch and saturation-loudness mappings used in the colour encodings were based on two prior findings on correspondences (Giannakis, 2001; Simpson et al., 1956). Problematically, in Simpson et al.'s (1956) experiment on hue-pitch correspondences, luminance was not controlled for in the available selection of prototypical colours to participants. Since a prototypical yellow is brighter than a prototypical blue this finding may reflect luminance-pitch associations rather than hue-pitch (Spence, 2011). If this is the case, then hue-pitch mappings may not be an intuitive representation for participants, therefore more evidence is required as to whether a hue-pitch association is maintained without variations of luminance. Another factor is the influence of correspondences irrelevant to the encoding used by the participant, so luminance-pitch associations may exert an influence on users with a hue-pitch mapping. Any counter-intuitive mappings may discourage users from engaging with their mapping and the information it provides. This may have been further compounded by the lack of feedback during the task as to their performance, so without knowing their colour errors participants may not maintain their attention on the colour cues available. For the RGB-loudness mapping, if participants cannot remember the exact pitch-hue mappings, they can rely on overall loudness to indicate luminance, similar to the luminance-loudness condition which performed equally well at object discrimination. Unique to these encodings is that white outputs the most harmonic sound, which may run counter to potential saturation-harmonics correspondences (Ward et al., 2006). Future SSD mappings should try and avoid correspondences that can work at cross-purposes. While sinewave sounds were chosen for the present experiment, the lack of use of colour cues by participants may make more complex sound-colour correspondences such as vowel-colour correspondences a potential viable alternative for colour representation (Marks, 1975; Moos, Smith, Miller & Simmons, 2014).

4.5.2. Are all correspondences equally useful?

Surprisingly, for the greyscale approaches, presenting the same luminance information through either variations in pitch or loudness resulted in significantly different performance levels, especially with regard to identifying the overall luminance of an image. It appears as if even though correspondences for pitch-luminance and loudness-luminance are well established (Hubbard, 1996; Lewkowicz & Turkewitz, 1980; Ludwig, Adachi & Matsuzawa, 2011; Marks, 1974, 1987; Ward et al., 2006), not all correspondences translate equally well into information that can be effectively used in sensory substitution. One potential explanation for this is a lack of luminance understanding from the luminance-pitch mapping, however in the colour replication task there were no significant differences in luminance recreation between luminance-pitch and luminance-loudness mappings. Instead, the act of exploring over a natural image with constant variations in luminance might be easier to classify each luminance in turn as ‘brighter’ or ‘darker’ through loudness than pitch-changes. Another difficulty with using variations of pitch between 500 to 5000Hz is that pitch and subjective loudness interact, so that very low and high frequencies are perceived as quieter (ISO, 2003). For the luminance-pitch mapping, the perceptual change in loudness may be misread as a meaningful cue or distract from concentrating on pitch alone. This is the first demonstration that not all visual-auditory mappings based on correspondences work equally well in communicating visual information in sensory substitution. This evidence brings up the possibility that specific types of correspondence may be more suitable for quickly and effectively communicating visual information, for example, if luminance-loudness correspondences are based on intensity-matching between the auditory and visual cortices (Goodyear & Menon, 1998; Jäncke et al., 1998), then this might be more efficient than other types of correspondences based on other factors relating to brain structure, environmental learning or higher-level associations (Spence, 2011; Walsh, 2003). Future studies on sensory substitution should therefore not take the use of any correspondences as a necessarily optimal pairing, instead each pairing needs to be separately examined to see if they are suitable for use in SSDs.

4.5.3. Optimal representations for different tasks

In the colour replication task, it was found that the hue-pitch mapping had significantly less error in human perceptual colour space than the RGB-loudness mapping. It was predicted that the independent dimensions for hue, saturation and luminance used in the hue-pitch mapping may be easier to decode back into a single colour than the related dimensions found in the RGB-loudness mapping. Interestingly the improved colour discrimination found from the hue-pitch mapping did not lead to increased scores for object discrimination or a reduction in object-based errors either. So the hue-pitch participants did not seem to effectively translate their colour knowledge to the colour variations found in natural images. Instead the hue-pitch mapping had the highest level of errors in identifying the overall luminance in an image which also held back overall object discrimination scores. It appears that participants do not seem to integrate the vibrotactile feedback representing luminance information effectively during naturalistic tasks. Potentially participants may need to concentrate on integrating these which may make this representation more optimal for smaller amounts of colour discrimination rather than the continuous variations present in naturalistic stimuli. As such, the discrimination of simple computerised colours may not be necessarily informative of their application to more practical natural colours useful for visual rehabilitation using SSDs (Maidenbaum, Abboud & Amedi, 2014; Reich, Maidenbaum & Amedi, 2012). Relevant to visual rehabilitation, in section 4.3.3 we show some exploratory work into replicating complex scenes or images with the colour encodings demonstrating that it is viable to use these to build up a decent understanding of a complex image through variations in colour and location (see fig. 4.11).

4.5.4. Future directions

The colour representations explored in this experiment are based on orthogonal colour spaces typically used in colour pickers (HSL) and those used in computer models of colour (RGB). While many previous colour SSDs have used variations of these colour spaces (Hamilton-Fletcher & Ward, 2013), their selection is largely arbitrary. A potential way to make exploration of colour more intuitive is through the use of colour spaces based on human colour perception such as CIE LUV space (Fairchild, 1998). This colour space has several unique points that lend themselves well to colour sensory substitution. The first aspect is the use of focal hues as endpoints of each dimension, L* indicates black and white, U* indicates red and green while V* indicates blue and yellow. From this six focal colours can be derived which can be linked to correspondences for representation. Colour opponency is also present here, with colours which inhibit their opponent in veridical colour perception positioned as opposites on each dimension. This large distance between opponent colours can translate to perceptual dissimilarity in their auditory or tactile representation. The closest parallel with this is the use of CIE LCH space (CIE LUV described with polar co-ordinates) as a basis for colour centroids in the ColEnViSon SSD (Ancuti et al., 2009). However this SSD segments this colour space into ten discrete colours reducing its colour resolution and represents these using an arbitrary colour-sound mapping. Human colour spaces such as CIE LUV space could provide an intuitive high resolution colour space and suitable correspondences for its dimensions may further enhance its usability in the sensory substitution of colour.

4.5.5. Conclusion

In conclusion, the evidence presented here may provide important considerations for representations of colour information in SSDs. The first consideration is that the additional dimensions of colour do not necessarily translate into improved performance in naturalistic discrimination tasks. Part of this can be seen in the equal amount of object errors seen across all mappings to which colour is a meaningful clue. It is therefore suggested that a lack of engagement with colour cues led to equivalent performance between the best performing greyscale and colour approaches. The most influential factor to performance on naturalistic tasks was the type of representation, namely mappings that used luminance-loudness correspondences performed best, irrespective of whether hue or saturation information was available. Interestingly, the presentation of the same luminance information using pitch did not perform equally well in naturalistic tasks despite also being a correspondence and showing an equal understanding of luminance information in abstract colour tasks. Therefore it appears that not all correspondences are equally useful with regards to sensory substitution use. Future studies should look to find optimal colour representations to increase engagement with colour cues in order to reduce errors potentially avoided by colour. For the colour representations it was also found that tasks involving abstract computerised colours did not correspond to higher scores in naturalistic tasks. As a result, different representations of colour may be optimal for different tasks and that abstract colour tasks may not be indicative of performance in ‘real world’ scenarios.

5. Sound-Colour Correspondences: Effects of Pitch, Noise, and Formant Structure on Associations to Equiluminant Colours

5.1. Abstract

Is the sounding of /u/ more likely associated with red than green? If so, why might this be? Recent studies have reported sound-colour correspondences between saturation and timbre, as well as between hues and vowel sounds. These findings have prompted conclusions that the structural properties of these sounds are driving these associations. However, since these colour selections do not control for variations in luminance, these findings could still reflect pitch-luminance associations between focal colours that also vary in luminance (e.g. focal yellows are more luminant than blues). We sought to replicate these associations in equiluminant conditions and further explore how variations in sound structure affected colour selections. In the first two experiments, participants were asked to choose which equiluminant colours 'went best' with specific sounds. In our first experiment, we sought to examine if vowel and complex sine wave sounds that share the same formant structure, but differ in their richness and recognisability would share the same colour correspondences. We report a 'red-green' correlation with both the formant ratio and high frequency content in vowel sounds, however for complex sine wave sounds only a new 'blue-yellow' correlation emerges with the amount of high frequency content. This indicates that formant structure is only one of the factors behind these colour selections. In addition, explorations of timbral sounds revealed that harmonics drive increases in colour saturation irrespective of hue. In our second experiment, we explore colour correspondences of simple auditory dimensions (pitch, loudness), their influence in complex vocal sounds and the influence of context on both. We find loudness-saturation and pitch-saturation correspondences occur in different pitch-ranges, indicating that relatively 'high' and 'low' pitches drive saturation changes rather than specific frequencies. In contrast to this, saturation towards particular hues seem to be strongly influenced by specific frequencies in both sine wave and vocal sounds. Articulation of vowels by different speakers led participants to re-apply previously established correspondences across all speakers simultaneously, indicating the flexibility of this link and the importance of context. In our third experiment, we tested the hypothesis that sound-colour correspondences are maintained under the influence of LSD. We

found that pitch-luminance, noise-luminance and harmonics-saturation correlations occurred in both LSD and non-LSD conditions and that there was a tendency towards previously observed vowel-colour correlations.

5.2. Introduction

In the opening moments of Disney's 1940 animated production of Fantasia, an abstract visual interpretation of the sounds of the Philadelphia Orchestra playing Bach's staccato fugue is created for the viewer. The abstract images vary in colour, shape and location and are the results of visual artists' mental imagery. The full bodied steady brass section gives rise to all encompassing swathes of deeply saturated reds and oranges, while the high pitched string instruments give rise to smaller angular columns of bright light with rapid changes in notes reflecting fast motion across the screen, tending to rise with ascending pitch or fall with descending.

In experimental studies, variations of luminance, location, shape, and saturation serve to reflect changes in pitch, loudness, tempo and timbral qualities (Marks, 1974; Walker, 2012; Walker, Francis & Walker, 2010; Ward, Huckstep & Tsakanikos, 2006). These intuitive pairings across the senses are referred to as cross-modal correspondences in the general population and congruency effects from this appear to influence their aesthetic appeal, integration and perceptual processing (Spence, 2011; Ward, Moore, Thompson-Lake, Salih & Beck, 2008). Correspondences are often contrasted with developmental synaesthesia, where stimulation in one modality (e.g. auditory) can elicit automatic, consistent and conscious percepts in a second modality (e.g. vision) in a small portion of the population (Novich, Cheng & Eagleman, 2011; Simner, 2012; Simner et al., 2006). It should be noted that correspondences and synaesthesia may share some of the same tendencies (e.g. pitch-luminance, Ward et al., 2006; for a review see Simner, 2013) and as a result it is likely that correspondences (especially those present in infancy) may influence any development of synaesthesia in the related modalities (Ludwig & Simner, 2013; Simner & Ludwig, 2012; Walker et al., 2010). In cases of visual deprivation, cases of acquired synaesthesia seem to be most commonly manifested as auditory to visual synaesthesias, suggesting strong predispositions for these regions to connect (Afra, Funke & Matsuo, 2009). As a result, auditory stimulation seems ideally suited for transferring 'visual' information. In sensory substitution devices, visual parameters are systematically encoded in sound with visually-deprived long term users of such devices reporting a 'visual' phenomenology from sound (Ward & Meijer, 2010).

Examining these multi-modal interactions allow us to identify rulesets that the brain uses to pair seemingly separate stimuli together. By examining the nature of these mappings, it becomes possible to determine which neural processes are likely to facilitate such bindings; such as whether these are driven by innate lower-level features and encoding (e.g. increasing loudness and increasing luminance, Marks, 1987) both of which are encoded via increased neural activation in primary but not secondary cortices (Goodyear & Menon, 1997; Mulert et al., 2005); through learned association (e.g. high-pitched and small objects - Evans & Treisman, 2010) where smaller animals tend have higher pitched voices (Fitch, 1997); higher-level cognitions (e.g. *High* pitch and *high* elevation, sharing the linguistic term 'high') including mediating pathways such as emotional valence, where emotionally-positive sounds and colours become associated (Palmer, Schloss, Xu & Prado-León, 2013; Sebba, 1991). The level at which a correspondence occurs has specific psychological implications, whereby lower-level correspondences are more likely be present in early brain development and affect lower-level perceptual tasks (i.e. speeded-classification tasks), while higher-level correspondences are likely to only affect untimed cross-modal matching tasks. Spence (2011) provides three classifications of correspondence which describe a particular correspondence's origin and influence on cognition. From this, the cognitive and neurological mechanisms that give rise to these correspondences can be better understood. The three classifications are as follows:

- Structural – An innate correspondence (subject to age of brain development) that affects lower-level perceptual and higher-level cognitive tasks.
- Statistical – A learned correspondence resulting from continual exposure to matching stimuli. This affects lower-level perceptual and higher-level cognitive tasks.
- Semantic – Learned as a result of a third mediating process, such as language development or emotion, where both stimuli may share a single word or emotion which combines the two. This is unlikely to affect speeded lower-level tasks however may still affect higher-level tasks.

While this provides a concise framework in which to examine correspondences, one must also be careful not to assume a correspondence only fits one of these descriptors. For example, while high-pitch and high-elevation might seem like a good example of a semantic correspondence, there is a preference for matching pitch and height in pre-verbal infants (Braaten, 1993), an effect of pitch-height on speeded classification tasks (Evans & Treisman, 2010), pitch-height correlations are also present in the environment as well as exaggerated by the structure of the ear (Parise, Knorre & Ernst, 2014), and finally pitch-height correspondences have a stronger effect in adults if this association is reinforced by language (Dolscheid, Shayan, Majid & Casasanto, 2013). Together these findings suggest a fundamental structural correspondence which is supported by our environment and anatomy before being semantically modulated by language.

The remainder of this introduction will consider two types of sounds: simple pure tones and complex timbral sounds. Each type of sound will be examined in their relationship to three properties of colour, namely luminance, saturation and hue. It will be argued that previous research has not yet convincingly disentangled these different properties of colour. Furthermore, it is suggested that complex timbral sounds are rarely broken down and analyzed in a meaningful way, making it unclear as to what auditory dimensions drive these colour changes.

5.2.1. Pure tone correspondences

The structurally simplest form of sound is a pure tone sine wave, which consists of a frequency (heard as 'pitch') and amplitude (perceived as 'loudness'). However these two dimensions can interact, so that changing the frequency may change the loudness to the listener in accordance with an ISO curve (ISO, 2003). Varying these two parameters has revealed a wide variety of colour correspondences.

Increasing the pitch of pure tones increases the luminance of colours selected (Ward et al., 2006). This correspondence can affect perceptual processing (Marks, 1987; Martino & Marks, 1999; Melara, 1989) and is also shared by non-human primates (Ludwig, Adachi & Matsuzawa, 2011). The occurrence of the same direction for pitch-luminance affecting children, related primates, adults, higher-level and lower-level tasks suggests this correspondence occurs innately at the structural level. Potentially this may mean that stimulus features are being cortically represented in similar fashions, or is related to the tonotopical organisation of the auditory cortex. The finding that pitch-luminance associations are experienced consciously in sound-colour synaesthetes, means that the pitch-luminance correspondence is a likely influencing factor in the development of specific synaesthetic mappings (Thornley-Head, 2006; Ward et al., 2006). Since pitch reflects the sole frequency in pure tones, these colour associations may also manifest in the additional frequencies found in more complex auditory stimuli.

Similarly increasing loudness also relates to increased luminance in both children and adults (Bond & Stevens, 1969; Marks, 1987; Root & Ross, 1965; Stevens & Marks, 1965). However it is not quite as uniform as pitch-luminance mappings, as for a smaller subsample of individuals, increasing loudness can be more intuitively put with *decreasing* luminance (Marks, 1974), however this may be influenced by the frequency of the tone played.

Likewise, other dimensions of colour have a more complex or less reliable relationship to pure tone stimuli. Ward et al. (2006) observed an inverted U-shaped relationship between saturation and increasing pitch, whereby pitches near middle C (262Hz) would derive the most saturated colours. However it is likely that this also reflects the pitch-luminance association as it is impossible for extremely dark or bright colours to also be highly saturated whereas colour with a moderate luminance can potentially be highly saturated. So one would expect to see this distribution if participants chose a random selection of saturations and hues, yet still maintained their pitch-luminance tendencies. Giannakis (2001) reports loudness-saturation mapping for pure tones between 110-3520Hz, where louder sounds correspond to more saturated colours. This provides an alternative explanation for Ward et al.'s inverted U-distribution for pitch and saturation, as extremely high or low pitched pure tones can sound quieter to the listener as described in ISO curves (ISO, 2003).

Pitch-hue correspondences have been reported in children, placing the highest pitched with yellow/green, middle pitch with red/orange and lowest with blue/purple (Simpson, Quinn & Ausubel, 1956). Giannakis (2001) also found that non-synaesthetic participants would combine red/yellow with high pitches, green/cyans with medium and blue/magenta with low-pitches. In regards to sound-colour synaesthetes, Orlandatou (2012) describes an experiment finding that higher pitched sounds (both pure tone and complex) result in more saturated colours than low-pitched, and that these have a tendency towards yellowish hues. An important note in the direct comparison of these studies is the different frequency ranges used in each experiment, with Simpson's experiment ranging between 125-12000Hz stimuli, Giannakis' experiment is between 110-3520Hz and Orlandatou's is 50-3000Hz. The relatively stable 'high pitch to yellow hues' association may be influenced by the context the sound is presented in - so any sound that is 'high-pitched' in its current context, or that over a critical frequency pure tones become associated with yellow hues. Another factor to consider is the lack of controlling for luminance in all of these studies, which means that these hue associations might be driven by pitch-luminance associations and then picking prototypical colour exemplars, since yellow is the brightest focal colour (Spence, 2011).

The previous studies that have attempted to examine pitch-hue correspondences suffer from methodological flaws that make any accurate pitch-hue correspondences difficult to determine. The most prevalent problem is not constraining the influence of luminance (Giannakis, 2001; O'Hare, 1956; Simpson et al., 1956), the use of non-coloured corrected screens or difficulties in analysis, such as using a circular representation of hue (e.g. HSL or HSB colour space) where perceptually similar colours (e.g. purple and red) are mathematically far apart (i.e. Hue = .99 and .01 respectively), making comparison difficult (Ward et al., 2006). One potential way this can be addressed is turning polar into Cartesian co-ordinates, where saturation towards particular hues (red, green, blue, yellow) is analysed such as with CIE LUV or Lab colour space.

In support of pitch-hue correspondences being either structural or statistical in origin is a finding by Thornley Head (2006). He found that pitch-colour correspondences were not significantly mediated by higher-level interference (i.e. incorrect note naming) for either controls or synaesthetes, but that the distance between notes related to the distance travelled in RGB colour space. This suggests an ordered relationship between colour and pitch within individual subjects. However using RGB-space makes understanding how much of this is due to hue, saturation or luminance correspondences difficult to fully untangle. An interesting note for participants with perfect pitch is that they become less influenced by linear pitch-luminance increases, keeping more rigidly to colours selected for individual notes within an octave. As a result, strong sound-colour associations may play a more complex role when taking into account the level of musical expertise of synaesthetes or controls.

5.2.2. Complex timbral correspondences

Progressing from simple pure tone stimuli to richer timbral sounds has some important influences on colour selection. More complex auditory stimuli consist of a fundamental frequency which is the lowest frequency in a sound (perceived as 'pitch') and the addition of higher frequencies referred to as overtones (in music) or formants (in vowel pronunciations). Many factors can be considered when classifying these timbral sounds, such as the range between the fundamental frequency and highest frequency, the number of overtones/formants that occur within the range, the distribution of these overtones/formants (i.e. higher / lower frequency content, or compact / diffuse scattering of the frequencies), alongside whether the frequencies are harmonic or disharmonic. Finally this can be used to describe either timbre that does not vary over time (called static timbre, or monophthongs for vowels) or that do vary over time (called dynamic timbre, or diphthongs for vowels) typically in a attack-decay-sustain-release pattern as found for instrumental sounds.

Ward et al. (2006) found that timbral sounds such as piano were given more saturated colours in comparison to pure tone stimuli when being played on the same note for both controls and synaesthetes. The increased saturation of instruments in comparison to pure tone stimuli in particular could be explained through either an overall increase in frequencies that create a richer sound, or it may be due to adding harmonic frequencies that increase the pleasantness of instrumental sounds. Also in this context, luminances chosen for piano or string instruments were not statistically brighter than their pure tone counterparts which points to the fundamental frequency having a key role in influencing luminance. However in a different context, when ten timbres were compared all playing the same note, significant differences were found between instrument-type and their associated luminance / saturation. This means that luminance is also affected by frequencies beyond the fundamental and that some element of the timbral sounds varied the saturation content as well. Timbre-luminance ranged from the didgeridoo (darkest) to harp (brightest) while timbre-saturation ranged from the least saturated (didgeridoo / harp) to the most saturated (super tenor / guitar). As stated previously, minimum and maximum luminance has a constraining effect on the potential choice of saturation, so it is not surprising to see the most de-saturated colours occur with the harp and didgeridoo stimuli. The main shortcoming is the lack of spectral analysis on the instrument sounds themselves (which is further complicated by using dynamic timbre that varies over time) as such there is no apparent meaningful order of these instruments. However, work to classify timbres across multiple dimensions that are deemed the most distinctive to the listener has been attempted previously using dynamic and static timbre (Grey, 1975; Von Bismarck, 1974), as such these could provide the spectral phenomena most likely to influence colour selection. Semantic connotations may also have an effect, with harps in particular commonly associated with luminous religious symbols. Separating the influence of sound and semantics will prove invaluable in understanding how these correspondences operate.

Another area of interest for complex sounds and associated colours is that of vowel sounds in speech. In the production of vowels, air passing through the vocal folds creates a glottal wave, consisting of a pitch (the lowest frequency) and its harmonics which are multiples of the pitch. The pitch and its harmonics are influenced by the thickness of the vocal folds and length of the vocal tract. Since males have thicker focal folds and longer vocal tracts than females or adolescents they tend to have lower pitched vocals (Titze, 1994). The intensity of glottal wave's harmonics is influenced by the structure of the vocal tract and tongue positioning. Different tongue positions alter the loudness of specific bands of frequencies (see fig. 5.1). The spectral peaks of intensity at certain frequencies that make up and classify vowel sounds are referred to as formants (typically referred to as F1, F2 & F3). Variations in the distribution of these formants give rise to different vowel sounds commonly used in language (see fig. 5.2). As a result variations in pitch, formant positions or relationships can be used to establish colour correspondences to meaningful changes in complex sounds.

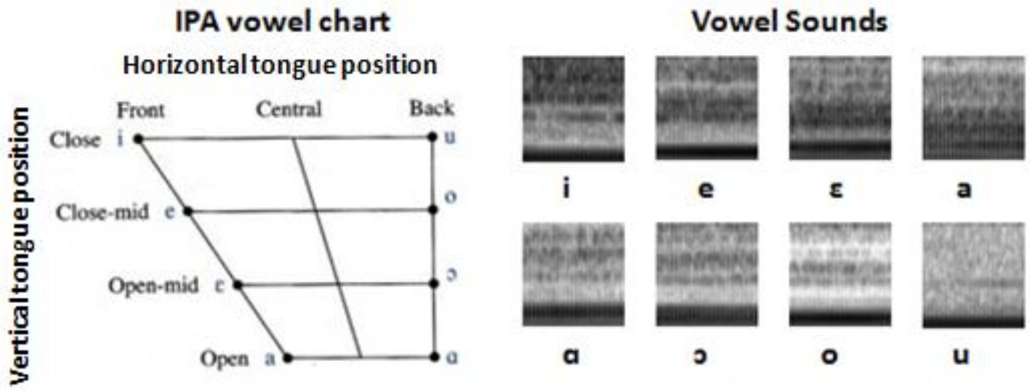


Figure 5.1. Left image shows an IPA vowel chart – the vertical position of the tongue changes vowels along the open-closed dimension, while the horizontal position of the tongue changes the front-back dimension, finally combinations of these produce specific vowels. The Right image shows the spectrogram results of these vowel sounds for a typical male speaker, vertical axis indicates frequency, horizontal axis indicates time and darker greys represent increases in amplitude.

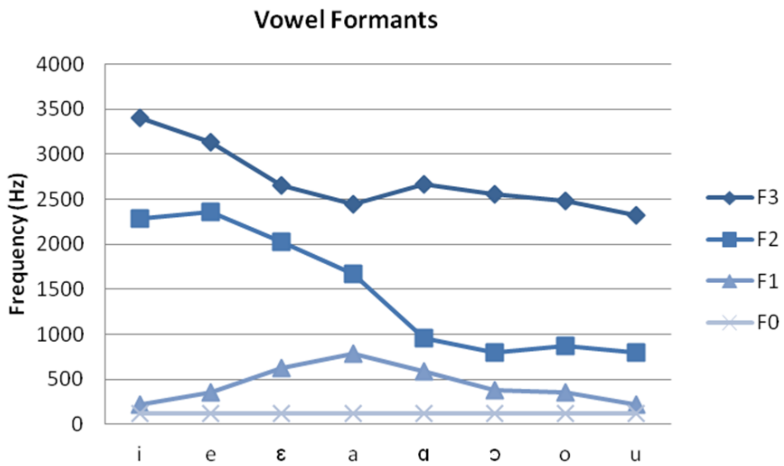


Figure 5.2. Pitch (F0) and formant frequencies (F1, F2, F3) across the main vowel sounds, specific frequencies taken from Moos et al. (2014).

In a meta-analysis of 38 studies on spoken vowel-colour synaesthetes, Marks (1975) standardized colour labels reported as proportions along Red-Green, Blue-Yellow and Black-White dimensions (see table 5.1). Marks concluded that red-green variance could be best accounted for by the ratio of formants 1 and 2 (excluding /i/), whereby having these two formants close (or 'compact') increased the redness of colours chosen, while having them far apart (or 'diffuse') increased the greenness of colours chosen. Subjective 'auditory brightness,' defined as the extent to which sounds can be described as 'bright' or 'dull' was also speculated to potentially drive yellow-blue colour dimensions. Marks noted that subjective 'brightness' increases with higher frequency content present in a sound, whether this is from the pitch or frequencies above that in a complex sound. Marks predicted that high and low subjective 'auditory brightness' would contribute to increased yellowness and blueness respectively in accordance with Simpson et al. (1956; and Orlandatou, 2012). For his exploratory analysis, Marks tentatively approximated this quality to the second formant frequency (F2), but noted this was only due to the absence of better auditory descriptors. Between F2 and yellow/blue saturation, only a small tendency in this direction was found. Since there is a lack of agreed upon objective auditory characteristics describing 'auditory brightness,' going forward, auditory brightness will be understood in terms of the sum of the pitch and the formants of vowel stimuli, or for stimuli without formants, the average frequency as weighted by power which is also known as a sound's centre of gravity. Both of these approaches give a proxy of 'auditory brightness' by having both pitch and the addition of higher frequencies influence its value, but are not exclusive to them. For vowel-colour associations, Marks' prediction would be that /a/ would be red, /e/ green, /i/ yellow & /u/ blue. The meta-analysis primarily consisted of French and German synaesthetes so potential semantic mediations include /ø/ for bleu, /au/ for blau, /ɛ/ for vert, /y/ for grün, /u/ for rouge, /o/ for rot, /o/ for jaune and /ɛ/ for gelb. The only potential semantic mediations are /o/ for red in German synaesthetes and /ou/ for red in French synaesthetes (see table 5.1), it could also be argued that /ɛ/ for yellow with German synaesthetes is perceptually close to /e/ (see figure 5.2), which accounted for the highest proportion of yellow responses. The highest proportions (i.e. /a/ and redness) however are not easily accounted for by semantic mediations. An alternative explanation of these relationships includes statistical factors, with the most common exemplars of vowels and colours becoming automatically associated (e.g. the written ‘a’ is more commonly associated with ‘red’ in synaesthetes and nonsynaesthetes – Simner et al., 2005). While semantic factors do not seem to be the most important factor, it is still unclear whether statistical or structural reasons best account for these findings, as both are suspected to affect speeded classification tasks (Spence, 2011) another approach is needed to disentangle these explanations.

Synesthetic Colors Induced by Vowel Sounds, Presented as Scores on Three Opponent-Color Dimensions

Vowel	Blue–yellow	Red–green	White–black	N
a	39.62 blue	111.88 red	34.75 black	419
e	54.5 yellow	11.5 green	49.5 white	400
i	47.75 yellow	44.75 red	67.5 white	400
o	44.5 yellow	93.25 red	54.5 black	372
u	36.75 blue	8.0 green	84.0 black	362
u_{French}	18.75 blue	31.0 green	23.5 black	195
u_{German}	14.25 blue	20.25 red	44.25 black	146
ou	0.37 yellow	21.88 red	35.75 black	157

Note. Data derived from Fechner (1876), Bleuler and Lehmann (1881), Flournoy (1892, 1893), and of 35 small-scale studies enumerated in text.

Table 5.1. Frequencies of colours chosen by synaesthetes for vowels along three dimensions of colour (Table from Marks, 1975).

Marks' finding has been supported and expanded to Polish and Japanese non-synaesthetes by Wrembel (2009) and Miyahara, Koda, Sekiguchi and Amemiya (2012) respectively with non-synaesthetes mapping English vowel sounds in a similar way to colours reported by Marks irrespective of their fluency in the language, once again suggesting that these correlations may not be semantically derived. However until now there had not been a formal spectral analysis of the sound given, it had been assumed that the pitch (or F0) remained relatively stable and that different speakers were comparable in the frequencies produced by the formant production, however phonetic aspects such as the gender of the speaker can have a marked impact on the F0 of vowel sounds, something that was found to effect the luminance of associated colours in a synaesthete (Fernay, Reby & Ward, 2012) and controls (Miyahara et al., 2006). Moos, Smith, Miller and Simmons (2014) experiment on vowel-colour associations in synaesthetes and controls sought to improve on several aspects of previous studies. Firstly, analyzing the vowel sound spectra of vowels from a trained phonetician in order to standardize the F0 component and identify the frequencies of the formants directly. Secondly, to correlate the F1 and F2 components to red-green, blue-yellow and luminance dimensions of participant's colour selections on a colour-corrected monitor in a perceptually accurate colour space (CIE LUV). Moos et al. found that the individual F1 and F2 components are correlated to redness-greenness of associated colours, such as higher F1 and lower F2s relate to increasing redness of colours, this correlation was also stronger in synaesthetes. Together these provide the strongest support so far for the notion of the compactness-diffuseness dimension being correlated to redness-greenness. No individual components appeared to correlate with the blue-yellow dimension (similar to Marks, 1975), however a 'brightness-dullness' dimension was not calculated for analysis. Separately, the F1 and F2 components both correlated to increased luminance when selecting achromatic stimuli. Taken in context with f0 also found to be related to luminance, this suggests that all frequency components are positively related to luminance, but to differing degrees.

5.2.3. Correspondences and visual photisms

Both correspondences and synaesthesia appear to map links between the senses in similar ways (e.g. high pitch and high luminance), previous researchers (Ward et al., 2006) have noted that these similarities in mappings may stem from a common underlying mechanism. Comparing correspondences in non-synaesthetes and synaesthesia is made more difficult by both structural and functional neurological differences (Rouw, Scholte & Colizoli, 2011). One possibility to examine if non-synaesthetes may be pre-disposed by correspondences to hallucinate photisms similar to those seen in synaesthesia is by examining temporarily acquired synaesthesia. While some previous attempts have been made through hypnosis, these experiments typically inform the participants to make specific hallucinations to given stimuli (Kosslyn, Thompson, Costantini-Ferrando, Alpert & Spiegel, 2010). An alternative that does not need to guide participants towards particular mappings is through transient acquired synaesthesia using drugs known to induce a similar visual phenomenology such as LSD through purely functional changes (Hartman & Hollister, 1963; Sinke et al., 2012). LSD has been reported to affect colour perception in a variety of ways. This includes reducing the ability to discriminate the hues of real colours and even increasing the persistance of flicker-induced colours, especially when in combination with pure tones (Abraham, 1982; Abraham & Wolf, 1988; Hartman & Hollister, 1963). As such, checks on the qualities of colours selected while under the influence of LSD may be important in determining if this can predict colours reflective of LSD hallucinations. Early evidence had indicated that LSD induced visual phenomena coincided with retinal changes suggesting a retinal origin of these disturbances (Krill, Wieland & Ostfeld, 1960). However, further evidence from totally blind users who did not have any retinal functioning but still experienced visual hallucinations from LSD, indicates that these hallucinations must involve cortical underpinnings (Krill, Alpert & Ostfeld, 1963). If the origin of LSD induced visual hallucinations are indeed cortical and can be elicited by auditory stimulation, to what extent is the content affected by widespread audio-visual cross-modal processing tendencies? LSD is known to produce a variety of spontaneous visual hallucinations, spanning from simple to complex geometric shapes, typically coloured and abstract in form typically spanning 7 to 10 hours (Sinke et al., 2012). Inducing these visual sensations from auditory stimuli only occurs in approximately 50% of LSD users, indicating large individual differences in the manifestiation of certain perceptual effects (Hartman & Hollister, 1963). Beyond visual disturbances, symptoms of LSD use also include difficulty in thinking and alterations of the user's emotional state (Isbell, 1959). These behavioural effects are supported by changes to neural activity of brain regions associated with decision making and emotional processing (Luke & Terhune, 2013; Sinke et al., 2012). Under the influence of LSD, will non-synaesthetes report any visual hallucinations that follow the sound-colour pairings seen in correspondences and developmental synaesthesia? If so, this would support the notion of correspondences providing a predisposition to specific visual hallucinations.

5.2.4. Experiments and hypotheses

In experiment 1 we explore whether previous vowel-colour findings can be entirely explained by their formant structure, for this we eliminated variations in the physical luminance of colours (as measured by a colorimeter) to get a purer measure of hue-saturation and compared vowels with sounds that share the same formant structure but lack their recognisability. Pastore, Li and Layer (1990) report that complex sine wave sounds generated from the formants that make up non-human chirps are organized in a perceptually similar way by listeners to the chirps themselves. As a result we decided to use vowel sounds and create complex sine wave sounds from combining pure tones at the frequencies found in vowels' F0-F3 components. Finally we will also further investigate Ward et al. (2006)'s finding that timbre is more saturated than pure tone, comparing harmonic to disharmonic will help answer if it is just due to more frequencies (i.e. noise), or whether harmonic content alone drives saturation. Firstly we predict that compact F1:F2 ratios will yield redder colours, while diffuse F1:F2 ratios will yield greener colours, for both vowels and complex sine waves. Secondly we predict that auditory 'brightness' (high frequency content) will be associated with yellower colours, while 'dullness' (low frequency content) will be associated with bluer colours for vowels and complex sine wave sounds. Finally we predict that harmonic content will be associated with more saturated colours than disharmonic content.

For experiment 2, while fundamental attributes of sound such as pitch and loudness have been repeatedly linked with luminance, their correlation with other colour variables is less well known. By testing participants in the same equiluminant conditions as experiment 1 we sought to examine pure tones' relationship with hue and saturation. We sought to examine how certain frequencies related to colour in a variety of contexts (100-3200Hz, 440-880Hz, vocal sounds) to see if previous pitch-colour relationships (Orlandatou, 2012; Simpson et al., 1956) could be replicated and which of these relationships were affected by context. We also sought to replicate previous loudness-saturation findings (Giannakis, 2001). We also examined the interaction between frequency and amplitude in sounds that spanned the same range by varying the 'auditory brightness' of vocal timbre or complex sine wave sounds. Finally previous vowel-colour relationships will be examined in a new context across multiple speakers, to see if vowel-colour relationships remain stable or are applied across the new context. We predict positive correlations between increases in saturation and more specifically yellow-saturation for increasing frequency in all variations in pure tone stimuli. We predict that increasing loudness of stimuli will increase saturation. For richer vocal and complex sine wave sounds we predict that 'auditory brightness' will correlate with yellow-saturation. For vowel-colour relationships we expect to see no influence of speaker on colour selections.

For experiment 3, sounds from our previous experiments (variations in pitch, harmonics and vowel sounds) will be given to participants who will be asked to select either their colour correspondences (while not on LSD) or the primary colour of any hallucinations (while on LSD). We predict that both correspondences and LSD hallucinations will not vary from one another. Based on previous experiments, we predict that increases in pitch will result in higher luminance (Marks, 1974), increases in noise will increase luminance (Lewkowicz & Turkewitz, 1980; Orlandatou, 2012) and vowel sounds will have increased red-saturation with more compact formants (Moos et al., 2014). In addition, we predict any correspondences found in experiment 1 and 2 to also be reflected here.

5.3. Method – Experiment 1

5.3.1. Participants

Twenty-five students at the University of Sussex participated. Their ages ranged from 18 to 31 (Mean 21.48, SD 2.58), 9 were female and 4 were left handed. Participants reported normal or corrected-to-normal colour vision and hearing, and actively denied the presence of sound-colour synaesthesia or any other synaesthesias.

5.3.2. Design

A repeated measures design was incorporated, with participants taking part in all conditions. Chroma from CIE LCH colour space was used for saturation dependent variable. The U* (red-green) and V* (yellow-blue) dimensions were taken from CIE LUV space as the U* and V* dependent variables. A high U* indicates increasing redness, a high V* score indicates increasing yellowness. The repeated measures variables were ordered according to their auditory characteristics, the harmonics stimuli were treated as separate conditions order from least to most harmonic. The vowel stimuli were ordered and weighted according to their scores on their two investigated auditory dimensions, namely compact-diffuseness & brightness-dullness.

5.3.3. Materials

5.3.3.1. Auditory stimuli

5.3.3.1.1. Vowel sounds

Eight natural vowel sounds were supplied by Moos et al. (2014). They are based on pronunciations of the primary cardinal vowels (i e ɛ a ɑ ɔ o u) of the International Phonetic Alphabet (International Phonetic Association, 1999) through recording a 65 year old male phonetician at 44.1kHz using a Sennheiser cardioid condenser microphone, Symetrix pre-amplifier, Edirol AD/DA converter and PC. Formants were identified using Praat (Boersma & Weenink, 2012) and confirmed using manual inspection of the frequency spectrum. Files were equalized in their loudness to 80dB using Praat and their length was constrained to 1 second using PSOLA in Praat. The final output was saved as individual wav files with 44.1kHz quality in mono.

5.3.3.1.2. Complex sine wave sounds

This auditory stimulation was created using Matlab to construct individual pure tones and were then combined to create the complex sine wave sounds heard by participants (each lasting 1 second). Eight complex sine wave sounds were constructed using the formant frequencies identified from the creation of the natural vowel sounds (see table 5.2). F0 and F3 were weighted at 50% amplitude, while F1 & F2 were weighted at 100% to avoid high frequencies perceptually dominating via increased loudness (ISO, 2003); this also improved their aesthetic appeal and emphasized their variation (since F0 is constant for all sounds).

For both the vowel sounds and the complex sine wave sounds the 'compact-diffuse' dimension (F1:F2 ratio) varied between /i/ at 10.17 and /ɑ/ at 1.62, while the 'bright-dull' dimension (F0+F1+F2+F3) varied between /i/ at 6035 and /u/ at 3460.

5.3.3.1.3. Harmonics

Stimuli for the harmonic-noise condition consisted of two main stimuli, the harmonic and disharmonic stimulus. The harmonic stimulus consisted of six sine waves made by combining 100, 200, 400, 800, 1600 and 3200Hz pure tones. Similar to the complex sine wave sounds, a 50% weight was given to the lowest and highest pure tones to reduce perceptual dominance of the highest pure tone. Frequencies between 200 and 1600Hz were given 100% weight. The noise stimulus consisted of bandpassed white noise between 100Hz and 3200Hz, so that the range remained consistent with the harmonic stimulus. The harmonic and noise stimuli were then paired together with the following weightings: 0-100, 20-80, 40-60, 60-40, 80-20 & 100-0. This provided 6 stimuli with varying emphasis on either the harmonic or noise based stimulation. All artificial sounds were sounded in Matlab and recorded as wav files at 44.1kHz quality using AVS audiorecorder.

5.3.3.2. Visual stimuli

Visual colours were displayed on a Dell D1626HT 20 inch CRT monitor using a 1024 by 768 pixel resolution with a refresh rate of 100Hz. An ATI radeon HD 2400 pro powers the monitor with a 32bit colour resolution. The red, green, blue and luminance output of the monitor (from 0 to 255) was measured using Cambridge Research Systems Lightscan v2 software with ColourCAL v2 colourimeter attached to the monitor in a darkened room. Sixty-four measurements were taken for each dimension, and a function was fitted to extrapolate these 64 points to the 255 potential measurements. A reverse function was modeled and provided the weighting of the R, G and B signals in a look-up table to normalize and linearise the RGB output accurately. By using the look-up table each increase in an R, G or B value in RGB space would increase the outputted luminance on the monitor by a predictable amount (see figure 1.8). The linearised RGB values allowed the calculation of CIE xyY values (R =0.627, 0.343, 11.601; G = 0.281, 0.615, 30.346; B = 0.151, 0.069, 4.21) which reflect the contribution of R, G and B channels to de-saturated luminance in the Y dimension, as well as hue and saturation chromaticity information in the x and y dimensions. CIE xyY space allows a transformation into CIE LCH and CIE LUV colour space (Fairchild, 2005; see section 1.4.9). In CIE LCH and LUV colour space, distances between colours reflect just-noticable colour differences, creating a perceptually uniform colour space. Both CIE LCH and LUV describe the same colour space and similarly separate out lightness information in their 'L' dimensions. Different from one another however is that they describe hue and chroma (saturation) information using either polar (LCH) or Cartesian (LUV) co-ordinates. It was deemed more intuitive for users to manipulate chroma and hue information independently (as seen in LCH models) rather than a mixture of both together (as with LUV) and so the chroma and hue dimensions were later used colour selection by participants in the experiment. As a stable luminance in CIE LCH space was utilised (65 of 100), the monitors output of maximum available chroma for various hues was assessed mathematically by charting the monitors estimated gamut range using CIE xyY values, it was calculated that maximum chroma available to participants was within acceptable boundaries for the monitors gamut, exceeding by less than 0.5% of the potential R, G or B value. The luminance output during the experimental task was also assessed to make sure that it remained stable during changes in chroma and hue by the participants; luminance varied less than 0.5% according to colourimeter luminance readings.

5.3.3.3. Stimulus presentation

Visual presentation was constructed using Matlab (The Mathworks, Inc.) with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Participants were presented with white instructions on a grey background of 60% luminance (in CIE LCH colour space). An initial training phase had participants listen to two randomly selected sounds from the experiment (excluding vowel sounds). One of these sounds was played again and participants were asked to alter the colour of a centrally presented circle to a colour that participants best felt matched the sound. This colour-selection process was then repeated for the second remaining sound. The coloured circle has an on-screen diameter of 1.75inches, taking up 5.01 degrees of the users visual angle at a typical viewing distance of 20 inches. The coloured circle was given a set luminance of 65%, set saturation of 20% and a random hue angle. Participants were not able to change the luminance, but could increase or decrease saturation (0 to 70%, so as not to exceed the gamut of the monitor) and cycle through the hue angle (0 to 360) in increments of 5 just noticeable differences to facilitate fast selection of colour. Participants could re-listen to the sound to be matched to it at any time, and were given unlimited time to make their matches before confirming their selection. Stimuli in each condition were presented in a random order. While the complex sine wave sounds and harmonic sound conditions were presented in a random order, the vowel sounds were always presented last. This was done to not influence the interpretation of complex sine wave sounds as similar to vowel sounds. Participants would listen to each sound in a condition first, before being asked to match each individual sound to a colour, so that context for each sound could be provided. Sounds were outputted using SoundMAX HD audio ESP and heard through HD 497 Sennheiser headphones. Participants’ starting colour, amount of increments, time taken and final selected colour was recorded in CIE LCH space. For analysis, polar co-ordinates from CIE LCH space is converted to Cartesian co-ordinates and rotated 12 degrees to create CIE LUV space where prototypical red is orientated at 0 degrees, in line with the CIE U* dimension (Boronkay, 2010). CIE U* and V* values were used as the dependent variables for multi-level modelling. Typical time taken was 15 minutes for participants.

5.3.4. Procedure

Participants were primarily recruited online using the University of Sussex’s Sona program, participants filled out a pre-screening form that involved asking about experiences of synaesthesia or auditory and/or visual impairments, to which all participants responded negatively. Prior to each experiment, the CRT monitor was turned on for 30 minutes to stablise the colour output to previously calculated levels from the colourimeter testing. On the day of testing, participants were presented with an information sheet, consent form and demographics form. Participants were asked about their colour vision and auditory perception, as well as any potential synaesthetic perception. Then participants were sat at the testing computer and were told they would be presented with a series of sounds, and then asked to match each sound in turn with ‘the colour that best matches or goes with’ that sound. They were walked through their controls for navigating colour space for the circle in the centre of the screen though changing hue, saturation, navigation speed, replaying the sound and finally confirming their selection. Participants were given two stimuli to practice on, and when they were ready to begin were left in a darkened room to concentrate on the task. The task takes 25 minutes to complete and was either paid in course credits or £5 depending on their preference, before being debriefed on the nature of the experiment.

5.4. Results

First participants' scores for individual conditions that had both a very low amount of time taken (mean <2 seconds) and low distance travelled from their randomly allocated starting point (mean <30 units) were excluded as outliers. This criterion was used to exclude participants who routinely accepted the first randomly presented colour to them and were presumably not engaging with the task. Hypotheses 1 and 2 were analysed using linear multi-level modeling, this was used to combine the standard analysis of a repeated measures ANOVA with a scale ordering of the auditory stimuli as used in a regression. This approach means that stimuli that are similar on an auditory dimension are expected to vary less in colour selection than with those that differ greatly on that dimension. This is used over a standard repeated measures ANOVA that treats all stimuli as equally separate (Hoffman & Rovine, 2007). The independent variables were weighted according to the IV's auditory dimensions of compact-diffuse (F1:F2 ratio) or bright-dull (F0+F1+F2+F3) fitting a linear model to the CIE U* and V* dimensions. Due to the frequency being logarithmically perceived, auditory dimensions were square-rooted so that the compact-diffuse and bright-dull dimensions were not exaggerated in their expected perceptual influence. The relevant auditory dimension was put as a fixed effect and as an unstructured covariate, participants were put as a random effect and parameter estimation was done according to maximum likelihood.

5.4.1. Vowel sounds

As the compact-diffuse and bright-dull dimensions are calculated using different dimensions, they were transformed for comparison between the two across the eight stimuli. This revealed that as auditory stimuli got increasingly compact (by an 8^{th} of the total range), CIE U* values would get redder by 1.76 units (r^2 = .031), in addition, as stimuli were increasingly dull (by an 8^{th} of the total range) CIE U* values would also get increasingly redder by 1.48 units (r^2 = .028). Both of these correlations were significant, with t (168) = -2.58, p = .011 and t (168) = -2.48, p = .014 respectively (see fig. 5.3). This shows that the change in compact-diffuse and bright-dull dimensions had similar impacts on both the amount of redness-greenness and the strength of correlation. There was no significant effect of either the compact-diffuse or bright-dull dimensions on the CIE V* dimension, contrary to the expectations of Marks (1974) but not to the data of Moos et al. (2014). One important note is that in vowel sounds, the compact-diffuse and bright-dull dimensions are not entirely orthogonal and do correlate to some degree, however this alone does not explain the high correlations found here.

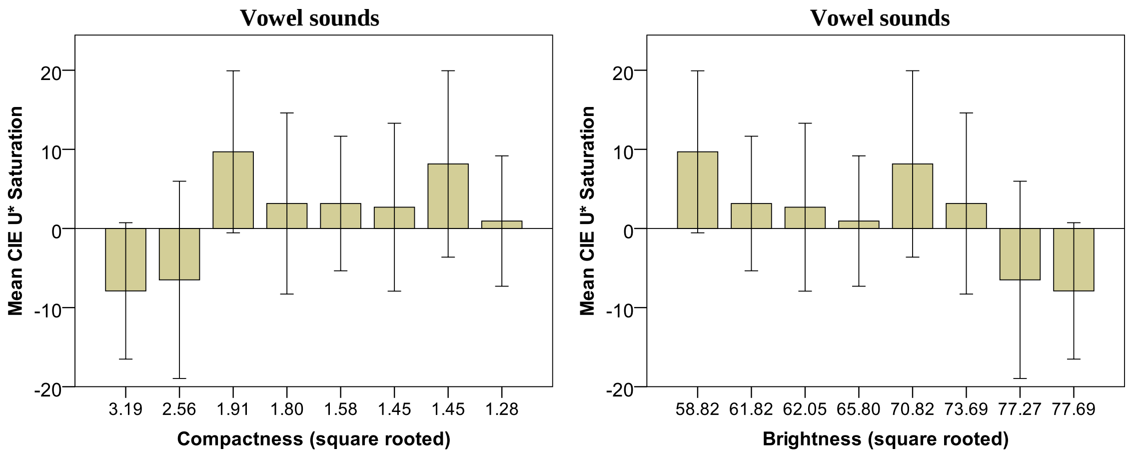


Figure 5.3. Vowel sounds. Auditory characteristics of compactness and brightness were found to correlate with CIE U* colour dimensions. The left graph shows the ordering of vowel sounds according to the compactness-diffuseness auditory dimension, high values indicate diffuseness while smaller values indicate compactness of the first two formants. Diffuse sounds yielded green colours (negative CIE U* values), while compact sounds yielded redder colours (positive CIE U* values). The Right graph orders the vowels in terms of their auditory brightness, with brighter sounds yielding greener colours while duller sounds yielded redder colours. Error bars indicate 95% confidence intervals.

5.4.2. Complex sine wave sounds

For complex sine wave sounds that lacked the recognisability of vowels but retained the structural relationship, a different set of correlations were found. First the significant CIE U* correlations found for vowel sounds were notably absent in the abstract stimuli, instead a new CIE V* correlation was found for the bright-dull dimension, 'brighter' stimuli (more high frequency content) would be associated with yellower colours (see figure 5.4), increasing the subsequent CIE V* dimension by 1.29 units per 8th of the bright-dull dimension that was travelled, and this tendency was significant, t (175) = 1.99, p = .048. This supports the findings of Simpson et al. (1956) and Orlandatou (2012) in that high frequency content is associated with increased yellowness in colours chosen, however here we see this correlation is independent from luminance.

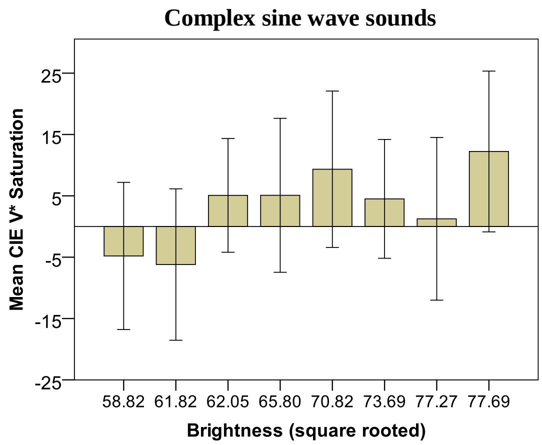


Figure 5.4. Complex sine wave sounds. A correlation between bright-dull auditory dimensions and the CIE V* visual dimension was found. It was found that 'brighter' auditory stimuli were associated with increased CIE V* values, yielding yellower colours, while 'duller' sounds yielded bluer colours. Error bars indicate 95% confidence intervals.

5.4.3. Harmonic sounds

To investigate whether noise or harmonic stimuli gave the most saturated colours, a repeated measures ANOVA was carried out for saturation in CIE LCH space. The results found that there was a significant main effect of harmonics on saturation, F (5, 120) = 22.969, p <.001, whereby more harmonic stimuli were associated with increased saturation, it was also found that a monotonic relationship best described this effect, F (1, 24) = 58.184, p <.001 (see fig 5.5).

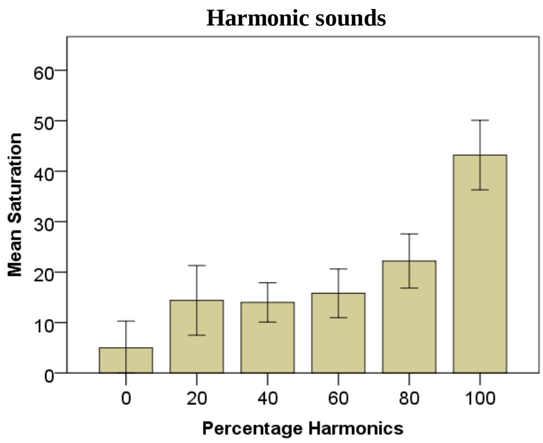


Figure 5.5. Noise-Harmonic sounds. Noise (left) transitioning to harmonics (right) is associated with increases in mean saturation for colours selected. Error bars indicate 95% confidence intervals.

5.5. Experiment 1 summary

Our results demonstrate a vowel-colour association where the auditory ‘compactness’ of the formants corresponds to ‘red-green’ colour dimensions as suggested by prior studies (Marks, 1975; Moos et al., 2014). This is the first direct correlation established between the two and further shows that this occurs independently of luminance. In addition to this, we also first report an equally strong correlation between the 'brightness-dullness' of vowels and the ‘greenness-redness’ of colours chosen. Contrary to prior expectations (Marks, 1975), the bright-dull dimension did not correlate with ‘yellow-blue’ dimensions. As such, previous findings suggesting a link between higher frequency formants and yellowness may primarily be the result of selecting higher luminances first, and then the nearest focal colour hues, such as yellow (Moos et al., 2014; Wrembel, 2009). Interestingly, when the pitch and formants of vowel sounds are abstracted to complex sine wave sounds, this results in an elimination the ’red-green’ correlation. Instead there is an emergence of a new correlation between the auditory brightness of stimuli and the yellow-blue colour dimension, with brighter stimuli yielding yellower colours. These findings closely follow previous findings where high pitch is placed with yellow hues for simple and complex sine wave sounds (Orlandatou, 2012; Simpson et al., 1956). However for the first time, it is shown that it that the yellow hue itself influences these selections independently of luminance. Differences in timbral quality between vowels and complex sounds derived from them may also explain these differences, such as the distribution of frequencies between the formant positions that add to the timbral richness of vocal sounds. Any influence of the distribution of non-formant frequencies remains to be determined.

Finally, it was found that the presence of harmonic frequencies drove increases in colour saturation. This suggests that it is the additional harmonics rather than just the additional frequencies (as seen in the disharmonic stimuli) that are likely the driving factor behind why instruments correspond to more saturated colours than pure tone frequencies (Ward et al., 2006). One explanation of harmonics being matched to increased colour saturation is through the auditory and visual stimuli being matched in terms of emotional valence (Palmer et al., 2013). A possible alternative explanation is through a learned association between disharmonics and desaturated colours as seen with analog TV static which emits both desaturated visual noise and auditory white noise.

5.6. Method – Experiment 2

While the first experiment provided new evidence on correspondences between sound and saturation / hue in equiluminant conditions, some questions remain unanswered. For vowel-colour correspondences we showed correlations between auditory brightness / compactness and ‘red-green’ colour dimensions. Are these correlations actually the result of participants applying concepts of ‘compactness’ and ‘brightness’ to the frequencies of the vowel-formants or is this the result of language, such that /i/ would always be associated with green irrespective of formant frequencies? To further explore these explanations we used vowel sounds from different speakers which varied in pitch and formant frequencies but placed these in the same context. If semantics is important, all speakers should show similar correlations, if the formant frequency relationships are important, it would be expected that ‘brightness’ and ‘compactness’ colour correlations are applied across all speakers. Experiment 1 also found that vowel sounds and complex sounds derived from them appear to use different colour dimensions, but why might this be? We sought to look at similar ranges of frequencies resulting from either sine waves or vocal sounds, and then varied their auditory brightness or their frequency bands to examine the influence these on colour selections. Finally, we explore fundamental auditory characteristics such as pitch, loudness and the context a sound is presented in equiluminant conditions to evaluate previous correspondences and elucidate relationships between colour and more complex sounds.

5.6.1. Participants

Forty-four participants were recruited from the University of Sussex, using the online recruitment service 'Sona.' Participants did not report any colour or auditory discrimination impairments or sound-colour synaesthesia. The participants age ranged between 17 to 37 years old (Mean 19.84, SD 3.10), thirty-three participants were female and three were left handed. Forty-three participants were Undergraduates while one was a post-graduate. Participants were either paid for their time or given course credits.

5.6.2. Design

Experiment 2 used the same repeated measures design as experiment 1, with participants taking part in all conditions. Likewise, the dependent variables were colour dimensions derived from CIE LCH and CIE LUV space, namely 'saturation' (CIE LCH's chroma dimension) and U* (red-green) and V* (yellow-blue) colour dimensions from CIE LUV space. The individual auditory stimuli that make up each condition were ordered from the lowest-pitch/loudness to highest-pitch/loudness when entered into the repeated measures ANOVA. The vowel stimuli were ordered according to the formant frequency distributions, namely compactness-diffuseness (ratio between formants 1 and 2) and brightness-dullness (addition of pitch and all formants together). This allowed a weighting of each auditory stimulus along the compactness and brightness dimensions for use in multi-level modelling.

5.6.3. Materials

5.6.3.1. Auditory stimuli

5.6.3.1.1. Pure tone frequency

As perceived pitch and loudness are related phenomena, loudness-equalised sine waves of varying frequency were produced. The amplitude of individual sine waves were scaled with reference to 40 phons (subjective measure of equal loudness) on equal-loudness-level contours (ISO, 2003). The frequency and amplitude of the stimuli are as follows; 100Hz (.92 amp), 200Hz (.68 amp), 400Hz (.52 amp), 800Hz (.4 amp), 1600Hz (.4 amp) and 3200Hz (.3 amp). This allows a variation in perceived pitch without obvious changes in subjective loudness to co-occur.

5.6.3.1.2. Pure tone frequency (within octave)

In order to gauge whether context is important, another set of frequencies were produced that spanned a shorter frequency-range. These consisted of sine wave frequencies taken from a musical octave, consisting of 440, 493, 523, 587, 659, 698, 784 and 880Hz pure tones. This range was chosen as it was close to the middle two frequencies in the previous condition and so would not consist of especially 'high' or 'low' frequencies relative to the previous stimulus. Since there are only minimal changes in loudness across these frequencies no loudness equalisation was applied.

5.6.3.1.3. Pure tone loudness

In order to create stimuli that varied in loudness but not pitch, a 40 phons 400Hz pure tone was created. Three additional stimuli were derived from this with 0.5, 0.25 and 0.1 amplitude proportions. This was done so that subjective loudness would be at normal levels with respect to the frequency condition, half-amplitude, quarter-amplitude and finally a 10th of the amplitude as the quietest stimuli.

5.6.3.1.4. Vocal frequency

In order to understand the impact of each frequency band on colour selections for vocal sounds, vocal stimuli without formants were created and then constrained to specific frequency bands. In order to create stimuli with the timbral quality of vocal harmonics of varying frequency range, an artificial vocal sound was created using Praat voice synthesis and analysis software. A 27 year old male vocal sound was recorded (F0 of 83Hz) and this was used as a reference by Praat for creating a synthesised vocal sound without formants. This allowed us to have a distinctive F0 with an equal distribution of power in all frequency ranges above F0 for a given time point while retaining the slight variations in pitch and loudness over time that typify the timbral quality of vocal sounds. This base sound was then bandpassed through either a 100-200Hz, 200-400Hz, 400-800Hz, 800-1600Hz or 1600-3200Hz gate. All sounds had the same implied vocal pitch since the louder frequencies are multiples of this F0. Each sound had the same loudness variations over time distinctive of vocal timbres and same power for each band. As a result, these stimuli allowed us to test individual frequency ranges for vocal-like sounds.

5.6.3.1.5. Complex sine wave brightness

In order to understand the influence of increasing the power of either higher or lower frequencies when the range of frequencies remains consistent, a complex sound was produced consisting of 100, 200, 400, 800, 1600 and 3200Hz sine waves together. This sound either had the lowest or highest three frequencies reduced in dB by 33 or 66%. Alongside the original sound this produced five sounds that vary in their power in low to high frequencies while retaining a 100 to 3200Hz distribution of frequencies.

5.6.3.1.6. Vocal brightness

An artificial vocal sound was created in Praat, using the same vocal reference and procedure as the 'vocal frequency' stimuli, however the base stimuli was bandpassed between 100 and 3200Hz. From this, frequencies either above or below 600Hz were reduced in dB by 33 or 66%. Including the original sound, this produced five sounds ranging between 100 to 3200Hz but varying in the power distribution of low (under 600Hz) or high (over 600Hz) frequencies.

5.6.3.1.7. Vowel sounds from different speakers

In order to examine whether the previously observed ‘compactness-red’ correspondence was the result of the recognition of the vowel or the phonetic sound of the vowel, the same eight vowels were created using speakers with varying pitch and formant positions. The different speakers were created in Praat using PSOLA voice re-synthesis software. Using the same vocal stimuli as the previous vowel-colour correspondence experiments, the eight primary cardinal vowels (i e ɛ a ɑ ɔ o u) of the International Phonetic Alphabet were pronounced by a 65-year old male phonetician. These sounds were pitch and loudness equalised and constrained to 1 second in length. The average formant frequencies were taken from across the vowel sounds (F1 to F4) in order to establish the vocal tract length of the male phonetician as 15.95cm, in order to synthesise an average male vocal tract length of 17.50cm, the pitch and formants were multiplied by 0.91 in Praat's change gender option to produce a prototypical male voice. To produce typical female pitch and formant frequencies (VTL of 14.60cm), a typical male pitch and formants would need to be multiplied by 1.2, in order to avoid resynthetising the male vocal stimuli twice, the original male vocal stimuli had its pitch and formants increased by 1.08 to produce the female pitch and formants. These procedures were used to create four sets of eight vowel stimuli, 'male pitch male formants,' 'male pitch female formants,' 'female pitch male formants' and 'female pitch female formants.'

5.6.3.2. Stimulus presentation

The same monitor, computer and setup were used as experiment 1 with one exception. The monitor was viewed through a one metre black tunnel designed to occlude any remaining residual luminance in the room from vision during the task. As a result of the longer viewing distances, the onscreen coloured circle (diameter 1.75 inches) now takes up 2.546 degrees of each participant's visual field.

5.6.4. Procedure

The procedure was unchanged from Experiment 1, with conditions played in a random order bar the vowel stimuli that were played last. The additional amount of stimuli also meant that the task took 25 minutes on average.

5.7. Results

5.7.1. Outliers

For each condition, the average time and colour-space travelled was taken for each participant. On average across all conditions the average time taken for colour selection was 8.33 seconds (SD 4.60), and the average distance in colour-space travelled in CIE LUV space was 46.08 units (SD 14.04). Participants who had conditions with both a very low time-taken (mean <2 seconds) and low distance travelled (mean <30 units) were considered as outliers. This resulted in the exclusion of one participant for the vowel sound condition. Greenhouse-Geisser and Huynh-Feldt corrected degrees of freedom are reported where appropriate according to Mauchly's test of sphericity.

5.7.2. Pure tone frequency

A repeated measures ANOVA found that increasing the frequency of loudness-equalised pure tone stimuli between 100 to 3200Hz is matched with increasing saturation, $F\ (3.55, 152.51) = 24.21$, $p < .001$ and that this trend is best described as linear $F\ (1, 43) = 57.67$, $p < .001$ (see fig. 6). Furthermore increasing frequency was also significantly related to increased red-saturation on the CIE U* dimension, $F\ (4.70, 202.95) = 4.96$, $p < .001$ with a linear trend $F\ (1, 43) = 12.83$, $p = .001$ (See fig. 5.6). Finally, frequency is also related to the yellow-saturation on the CIE V* dimension, $F\ (5, 215) = 3.83$, $p = .002$ and that this relationship was best described as cubic $F\ (1, 43) = 9.18$, $p = .004$ (see fig. 5.6).

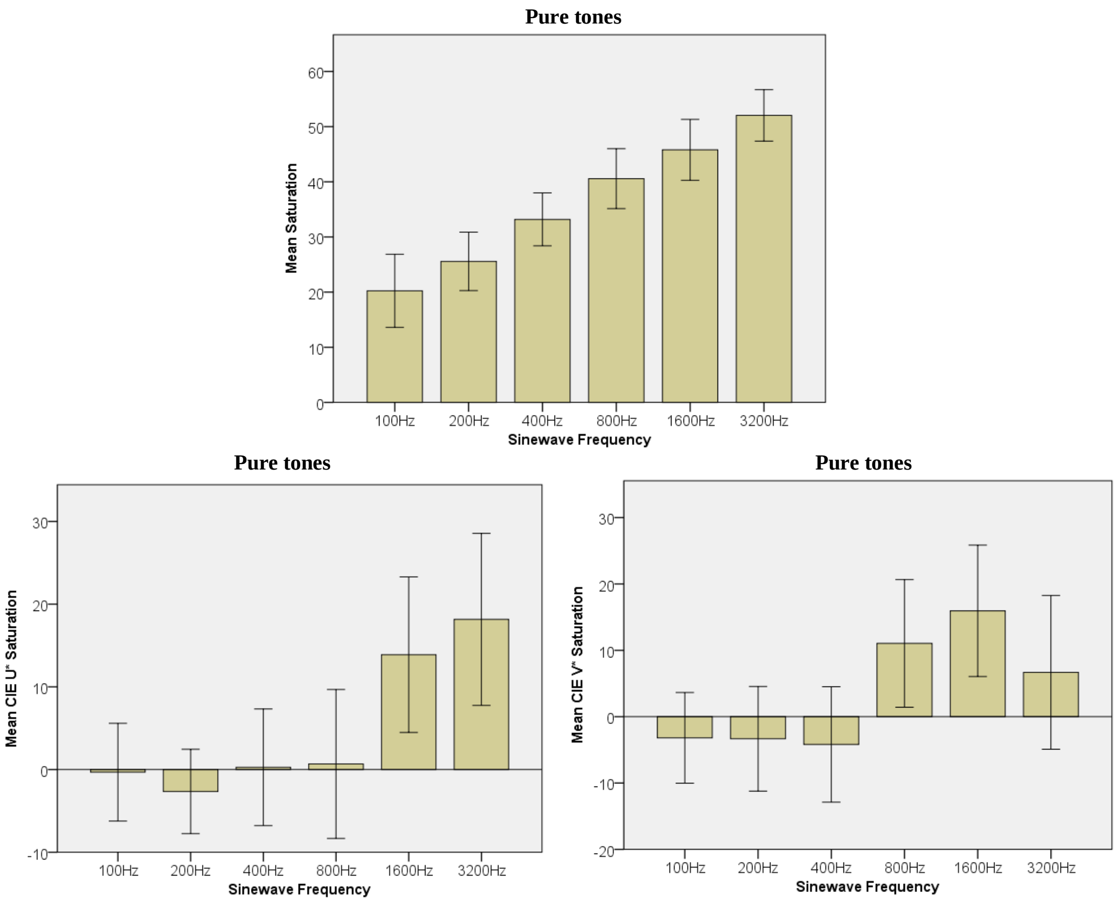


Figure 5.6. (top) Frequency to saturation graph, as frequency increased so did the amount of colour saturation chosen by participants. (left) Frequency to CIE U* saturation graph, as frequency increased, the amount of red-saturation increased as well. (Right) Frequency to CIE V* saturation graph, as frequency increased, the amount of yellow-saturation increased as well. Error bars indicate 95% confidence intervals.

Increasing the frequency while keeping the perceived loudness constant for pure tones appears to be strongly related to increased colour saturation, particularly towards hues with increased yellow and red components in equiluminant conditions. Pitch-saturation mappings have been previously reported in relation to both correspondences and synaesthesia as a cubic relationship with pitches near 262Hz being the most saturated (Ward et al., 2006). One explanation for the cubic relationship seen in Ward et al.'s experiment is that stronger pitch-luminance correspondences override any pitch-saturation mapping, and this appears to be the case here as controlling for luminance produces a highly significant linear pitch-saturation relationship.

5.7.3. Pure tone frequency (within octave)

While varying the frequency between 100 and 3200Hz provides highly significant colour correspondences, one aspect of this that requires further examination is the position of a frequency within a given context. This compares whether the absolute frequency is important for the previous correspondences or whether a frequency's relative position as 'low' or 'high' within a given context is what drives these correspondences. Therefore a new set of frequencies within a context of 440 to 880Hz were presented to participants. A repeated measures ANOVA found that increasing the frequency of these sine wave stimuli also increased the saturation of colours chosen by participants, F (6.47, 278.12) = 15.68, p < .001 and that this trend was linear F (1, 43) = 62.02, p < .001 (see fig. 5.7). Unlike previous pitch-colour findings, correlations with both the CIE U* and CIE V* dimensions were both non-significant with F (4.98, 214.23) = 1.67, p = .143 and F (6.45, 277.39) = 1.08, p = .373 respectively.

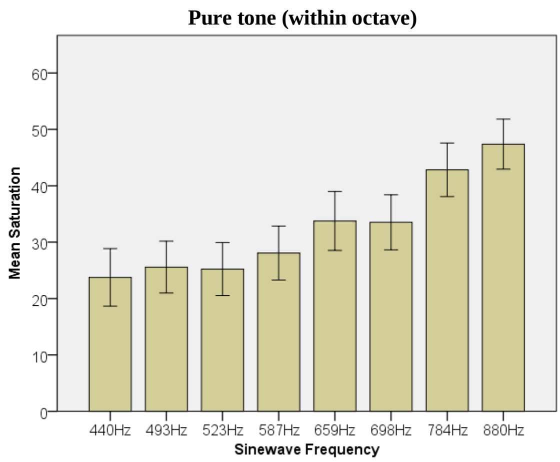


Figure 5.7. Frequency to saturation graph, as frequency increased within an octave so did the amount of colour saturation by participants. Error bars indicate 95% confidence intervals.

When providing a more constrained set of frequencies than the previous 100-3200Hz frequency condition, the pitch-saturation correspondence appears to reapply itself to its new context rather than maintaining a similar saturation level to 400 and 800Hz stimuli from the first condition. Of note is that there does not appear to be saturation towards particular hues. This is interesting as the prior pure tone frequency condition that spanned a larger frequency range uniquely had correlations between pitch and yellow-saturation as well as red-saturation. For the prior pure tone condition, red and yellow were particularly saturated above 1600Hz, notable since these are frequencies above the upper limit in this condition. This finding points to qualitatively different mechanisms used for saturation or hue mappings. For these pitch-saturation mappings to be flexible, their auditory characteristics need to be treated in a more abstracted sense, such as treating a frequency in terms of its magnitude as 'low' or 'high' pitched in a given context. By comparison, pitch-hue findings are more rigid, with participants refusing to put these new 'high' pitches with yellow/red hues, instead specific frequencies (above 800Hz) only seem to be able to elicit these correspondences. Previous studies finding pitch-yellow hue relationships have also utilised frequencies above 800Hz to find these (Orlandatou, 2012; Simpson et al., 1956).

5.7.4. Pure tone loudness

Another primary dimension in which sounds can vary is in their loudness. A 400Hz pure tone was played at either .10, .25, .50 or 1.00 Amplitude. A repeated measures ANOVA found that increased loudness was significantly related to saturation, F (2.22, 95.61) = 17.87, $p < .001$ and that this trend was linear in nature F (1, 43) = 28.04, $p < .001$ (see fig. 5.8). The ANOVA results found that neither CIE U* or CIE V* correlated with loudness, with F (3, 129) = .64, $p = .590$ and F (2.78, 119.42) = 2.12, $p = .106$ respectively.

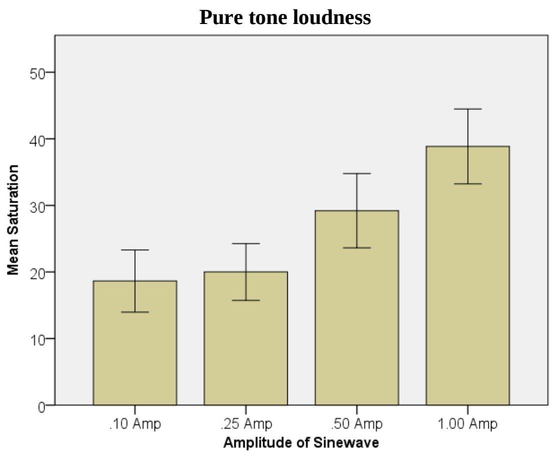


Figure 5.8. Loudness to saturation graph, as loudness increased so did the amount of colour saturation chosen by participants. Error bars indicate 95% confidence intervals.

The loudness-saturation finding mirrors a previous finding by Giannakis (2001) who found that increased volume of sounds correlated with increased saturation. Similar to a pitch-magnitude explanation of pitch-saturation correlations, a loudness-magnitude explanation would also account for positive correlations with increased saturation as well as brightness as seen previously (Lewkowicz & Turkewitz, 1980; Marks, 1974, 1987).

5.7.5. Vocal frequency

While pure tone frequencies have been well investigated previously, systematic changes in the frequencies of timbral stimuli have been less so. A repeated measures ANOVA found that increasing the pitch range of vocal sounds was significantly related to increasing the saturation of colours chosen, F (3.50, 150.51) = 6.05, p < .001 and that this trend was linear F (1, 43) = 12.30, p = .001. Increasing the vocal frequency was also significantly related to increasing CIE V* values, progressing from blue-saturation to yellow-saturation, F (3.45, 146.40) = 8.23, p < .001 with a linear trend F (1, 43) = 20.32, p < .001 (see fig. 5.9). However, unlike our pitch-hue findings, there was no significant correlation between vocal frequency and CIE U* dimension, F (4, 172) = .698, p = .595.

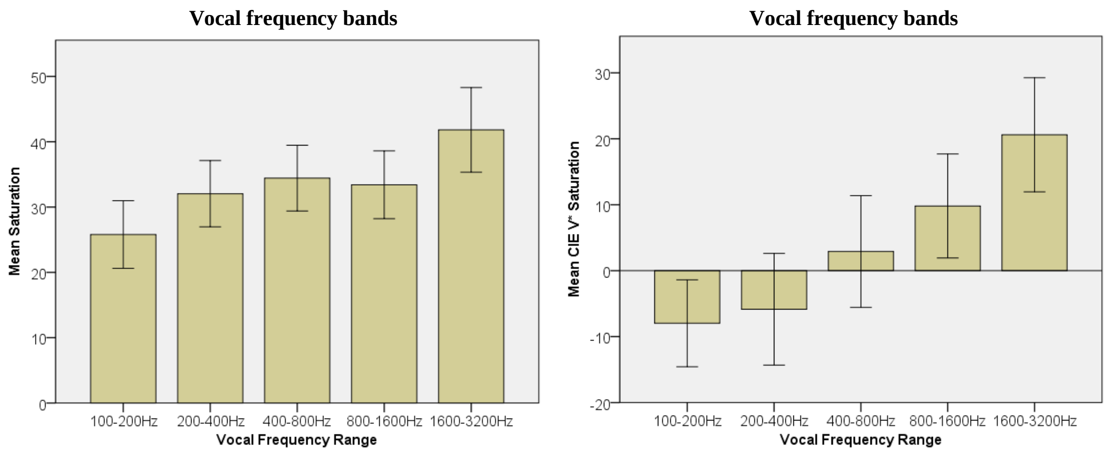


Figure 5.9. Vocal frequency to saturation graph, increasing the frequency range of vocal stimuli increased both overall saturation (left) as well as CIE V* saturation from blue to yellow (Right). Error bars indicate 95% confidence intervals.

Similar to pure tone frequency-saturation correspondences, higher frequencies in vocal stimuli also elicit saturation correspondences. Interestingly the type of hue was a very strong correspondence with vocal frequency range, with lower frequency vocal content associated with blue hues (negative CIE V* values) to higher frequency vocal content being with yellow hues (positive CIE V* values). This could explain some of the previous vowel-colour correspondences with vowels with higher frequency content being associated with yellow hues and lower frequency content with blue hues (Marks, 1975; Miyahara et al., 2012; Wrembel, 2009).

5.7.6. Complex sine wave brightness

These stimuli all have the same frequency range (100-3200Hz) however the loudness of either the bottom half or top half of frequencies are attenuated, so that the range remains constant but the centre of gravity (or 'brightness') of the sound varies. This can be described as the average frequency of a sound as weighted by its power. A repeated measures ANOVA found that increasing the brightness of sine wave stimuli also increased the saturation of selected colours F (2.86, 122.88) = 14.39, $p < .001$ with a linear trend, F (1, 43) = 30.62, $p < .001$ (see fig. 5.10). However, brightness was not associated with saturation towards any particular hues along the CIE U* or CIE V* dimensions, F (4, 172) = 2.33, $p = .058$ and F (4, 172) = .44, $p = .780$.

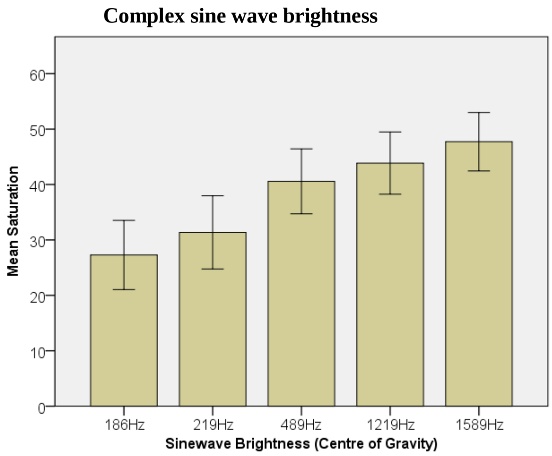


Figure 10. Sine wave brightness to saturation graph, increasing the centre of gravity (average frequency as weighted by power) increased overall saturation. Error bars indicate 95% confidence intervals.

Maintaining the same frequency range but changing the power of different bands of frequencies reveals a strong 'auditory brightness' to saturation correspondence as measured by centre of gravity values. It appears that despite the saturation varying, the hues chosen by participants are largely yellow (see fig. 5.11), however there was little to no variation of these strong CIE V* saturations to variations in auditory brightness. This indicates that the complex sine wave stimuli always had participants selecting colours with a yellow hue; however stimuli with a higher centre of gravity ('brightness') chose more saturated yellows. Taking previous pitch-yellow correspondences into account, it appears as though the presence of higher frequencies (even when attenuated) drives participants to choose yellow colours, albeit weighted by this frequency's loudness. These findings also help explain previous complex sine wave sound correspondences based on vowels, with high frequency content increasing yellow correspondences (experiment 1).

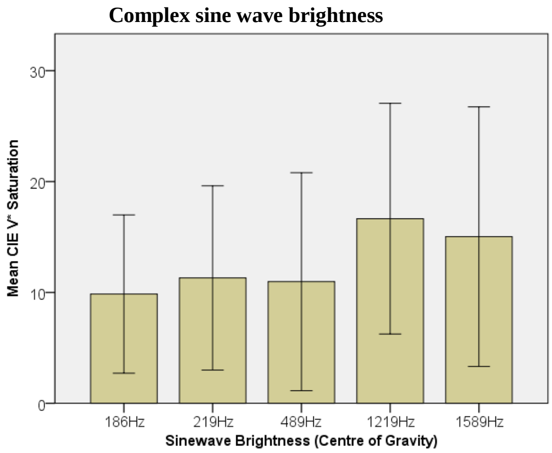


Figure 5.11. Sine wave brightness to CIE V* saturation graph, all variations of Sine wave Brightness stimuli had participants selecting yellow hues. Error bars indicate 95% confidence intervals.

5.7.7. Vocal brightness

Can the findings for sine wave brightness also be applied to stimuli with a vocal timbral quality? Vocal timbral stimuli ranging between 100-3200Hz with variations in 'auditory-brightness' were played to participants. A repeated measures ANOVA, revealed that changing the centre of gravity ('brightness') of vocal stimuli did not have a significant effect on saturation, CIE U* saturation or CIE V* saturation, with F (2.66, 114.27) = 1.03, p = .377, F (4, 172) = 1.01, p = .402, and F (4, 172) = .426, p = .790 respectively. However, similar to the sine wave stimuli, the vocal brightness stimuli also had strong yellow saturation (see fig. 12). So the presence of high frequencies as seen in the vocal frequency condition may indicate yellow saturation irrelevant of the distribution of the frequencies across the range.

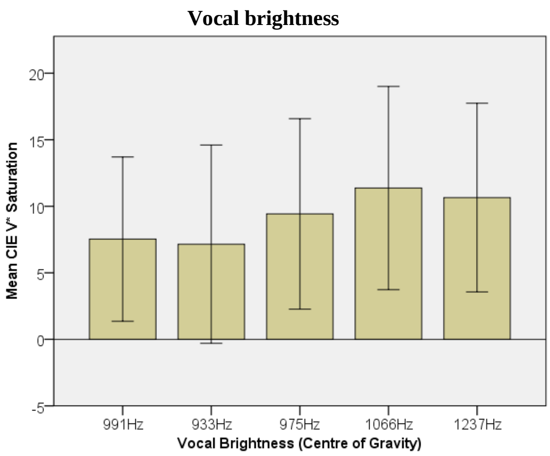


Figure 5.12. Vocal brightness to CIE V* saturation graph, all variations of Vocal Brightness stimuli had participants selecting yellow hues. Error bars indicate 95% confidence intervals.

5.7.8. Vowel sounds

Previous experiments have found consistent vowel-colour correspondences but only within the limited context of a single speaker across the eight cardinal vowels (i e ɛ a ɑ ɔ o u). In this context, experiment 1 found a ‘compactness-red’ correlation, however whether this is the result of semantics (e.g. /i/ relates to greenness) or phonetics (applying ‘compactness’ evaluations onto these formants) remains unexplored. To disentangle these explanations, we presented participants with four speaker types (varying in pitch and formant positions) pronouncing the same eight vowels all within the same context. A semantic explanation would predict each of the speaker types to have the same ‘compactness-red’ correlation across their eight vowels. If the recognition of the vowel drives these correspondences, then the specific frequencies of the formants should not matter. For instance, a /i/ sound would be expected to elicit green hues irrespective of the speaker. A phonetic explanation would predict that the frequencies of the formants are important and so a ‘compactness’ evaluation would be applied across the entire context and therefore all speaker types. For instance, in pronouncing /i/ a male speaker might have a compactness score (F1:F2 ratio) of 6.22, while a female speaker might have one of 5.66, a phonetic explanation would expect the male vowel to elicit a more saturated green than the female vowel.

For the multi-level modelling analysis "compactness" scores were square-rooted (as frequency is logarithmically perceived) and placed as a fixed variable, participants were random variables, the covariance structure was diagonal due to the repeated measures design and the model was calculated to 'maximum likelihood' for the CIE U* dependent variable. In order to test if semantics or phonetics best described our data, two separate analyses were run.

5.7.8.1. Vowels within speaker types

Interestingly, no variants of speaker managed to maintain previously observed compactness-red correlations, with t (180.51) = -1.20, p = .230 (Male Pitch, Male Formants), t (129.34) = 0.37, p = .710 (Male Pitch, Female Formants), t (171.92) = -1.52, p = .131 (Female Pitch, Male Formants), and t (150.84) = -0.05, p = .963 (Female Pitch, Female Formants). As such even previous correlations found for male speakers did not manifest within this new context. Taken in combination with the lack of a correspondence found for a prototypical male voice, it appears that the semantic aspect of each vowel does not inform their colour correspondences, so /i/ was not systematically paired with green across speaker types for instance.

5.7.8.2. Vowels across speaker types

One final explanation is that the change in context encouraged participants to apply a more flexible phonetic mechanism to all of the stimuli rather than individual speaker subsets. In order to explore this possibility, all speaker types were pooled together. This analysis found that compactness was marginally significantly related to CIE U* dimensions, t (633.78) = -1.75, p = .080, with more compact sounds associated with redder hues similar to previous studies. So it appears as though the change in context may have encouraged participants to extend the same auditory mechanism to the same colour dimensions but expanding to all stimuli.

5.7.8.3. Vowels – exploratory analyses

Further exploratory analyses across all speaker types yielded interesting effects that draw on some of our previous findings, with 'auditory brightness' positively correlated to a marginally significant degree with both saturation and CIE V* colour dimensions, with t (797.26) = 1.83, p = .068, and t (808.33) = 1.88, p = .060 respectively. This mirrors previous vocal frequency to saturation / yellow findings described in section 7.4 as well as predictions by Marks (1975). No other marginally significant correlations were observed, including the dullness-red correlation observed for vowels in experiment 1.

5.8. Experiment 2 summary

For pure tones our results demonstrated a strong frequency-saturation effect furthermore these correspondences adapted to fit a variety on contexts (i.e. 100-3200Hz or 440-880Hz), these all indicate an abstracted sense of magnitude (i.e. relatively 'low' or 'high') was responsible for this mapping. Previous investigations into pure tone-colour correspondences found strong frequency-luminance correspondences, with ‘middle’ frequencies being more saturated (Ward et al., 2006). Our evidence suggests that there *is* a linear frequency-saturation relationship, but that luminance correspondences may override these preferences when available. In vocal and complex sine wave sounds, added emphasis on higher frequencies also tended to increase saturation. Similarly for vowel stimuli, an exploratory analysis found that higher frequency formants increased colour saturation. In terms of specific hues, sounds that contained frequencies over 800Hz routinely contained yellow-saturations. Unlike the saturation mappings these hue correspondences do not appear to be affected by the context that a sound is in. Previous research on pitch-hue correspondences have found yellow to high pitch mappings (Orlandatou, 2012; Simpson et al., 1956), while these findings were potentially explainable through prototypical yellow's increased luminance component and pitch-luminance correspondences (Spence 2011), our research in equiluminant conditions suggests that yellow's hue component may also drive these correspondences. Since higher frequencies in vocal sounds also correspond to yellow hues, this may explain the previous auditory brightness to yellow saturations found in vowel stimuli and in our exploratory analyses across all speaker types (Marks, 1975; Miyahara et al., 2012; Wrembel, 2009). Finally there were a variety of interactions of loudness on saturation. For pure tones there was a simple linear relationship with increasing the loudness of a 400Hz tone leading to greater saturation. However for complex sine wave sounds that ranged between 100-3200Hz, only the increased loudness of the higher frequencies in the sound was correlated to saturation. Therefore, loudness is linearly related to saturation in sine wave sounds, unless it is specifically increasing the loudness of the lowest frequencies present in a sound. For vocal-timbre sounds, auditory brightness was not correlated in the 'vocal brightness' stimuli but was marginally significantly related to formant frequencies across vowels from multiple speakers. One possibility is that auditory brightness requires larger variations to elicit this correspondence in the 'vocal brightness' condition.

5.9. Method – Experiment 3

Correspondences share some commonalities with other cross-modal sensations such as those experienced in synaesthesia (Simner, 2013). It has been suggested that the common mappings between sound and colour dimensions may be the result of the influence of correspondences on the development of synaesthesia (Ward et al., 2006). If this is the case, correspondences may also influence other sound-colour mappings such as those reported to be experienced during temporally acquired hallucinatory visual sensations through drugs such as LSD (Luke & Terhune, 2013; Sinke et al., 2012). If these visual sensations are influenced by correspondences we would expect to see the correspondences to auditory stimuli seen previously may manifest themselves in visual disturbances. In the next experiment it is examined if sound-colour correspondences are maintained or used to inform visual hallucinations while under the influence of LSD.

5.9.1. Participants

Ten participants were recruited by Imperial College London, one was female, and the mean age was 34.20 (SD 7.40). Participants were eligible to participate if they had indicated they had previously taken LSD without adverse side effects requiring any form of treatment and had no personal history of mental illness or family history of schizophrenia. Participants were asked about previous drug use which featured ketamine, psilocybin and ayahuasca. Participants personal dose of LSD for the experiment varied between 40 and 80mcg as appropriate doses were determined through initial testing.

5.9.2. Design

A repeated measures design was incorporated with participants completing the sound-colour matching task over two days (day 1 - no LSD; day 2 - LSD) and for each day two times (time 1, 2). The colours chosen for each auditory stimulus were averaged between time 1 and 2 on each day, creating a single colour score for each stimulus on day 1 and 2 for analysis. Colour scores were converted from RGB into CIE LUV and LCH colour space which provided four colour dimensions for analysis, luminance (black-white), overall saturation, U* (red-green saturation) and V* (yellow-blue saturation). Each of the sets of auditory stimuli (variations in pitch, harmonics and vowel sounds) were ordered according to their most distinctive auditory characteristics, ordered using increases in pitch or harmonics for a repeated measures ANOVA analysis. For vowel sounds, 'compactness' and 'brightness' scores were worked out from their formant positioning for multi-level modelling analysis (see experiment 1 & 2).

5.9.3. Materials

5.9.3.1. Auditory stimuli

Stimuli were chosen that had strong sound-colour correlations from previous experiments. Pitch variations with equalised loudness were used (see experiment 2), consisting of 100, 200, 400, 800, 1600 and 3200Hz pure tones. These were chosen due to having been found to have a positive relationship between increased frequency and luminance, overall saturation, U*saturation and V* saturation (Marks, 1975; experiment 2). Variations in harmonics and disharmonics was also selected from experiment 1, with variations of 100% harmonics (simultaneous 100, 200, 400, 800, 1600 & 3200Hz tones) mixed with disharmonics (100-3200Hz bandpassed noise) to varying degrees (100, 80, 60, 40, 20 and 0% harmonics). Disharmonics has been found to increase the brightness of visual stimuli chosen (Lewkowicz & Turkewitz, 1980) while increased harmonics has been linked to increased colour saturation (Ward et al., 2006; experiment 1). For vocal sounds, vowels were chosen consisting of the primary eight cardinal vowels (i e ɛ a ɑ ɔ o u) pronounced by a 65 year old male phonetician as used in experiment 1 and Moos et al. (2014). Vowel sounds were classified according to their formant positions, either on the ratio of the first two formants giving a measure of "compactness-diffuseness" or the amount of high frequency content, measured by adding the pitch, F1, F2 and F3 together, for "brightness-dullness." All auditory stimuli were 1 second in length.

5.9.3.2. Visual stimuli

Participants were able to select colours using a colour swatch (see fig. 5.13). This was done to reduce time taken to select colours as seen in previous experiments. There were four versions of the swatch with both the colour order seen in normal or reversed, and the monochromatic options either at the top or bottom of the selection screen to avoid any systematic pitch-height correspondences influencing selection (Evans & Treisman, 2010).

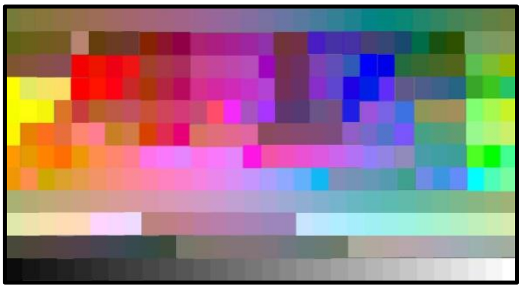


Figure 5.13. Colour swatch used by participants to make colour selections. Four variants of this were used to avoid spatial correspondences influencing colour selection.

5.9.3.3. Stimulus presentation

Visual and audio stimuli were presented using psychopy on a HP 9470m notebook on a 14inch screen. Auditory stimuli were presented to using participants using sandstorm SBS2112 sound system with the speakers placed in front of the participants. Participants were played the each auditory stimulus in a random order which was followed by a randomly selected colour swatch for participants to select a colour. During the non-LSD condition participants were asked to provide the most complementary colour to each sound played, for the LSD condition participants were asked to indicate the colour of any hallucination experienced as a result of the auditory stimulation. After participants chose a colour on the swatch, another screen asked participants to indicate the strength of any visual hallucination on the four-point perceptual awareness scale (Sandberg, Timmermans, Overgaard & Cleeremans, 2010).

5.9.4. Procedure

Prior to experimentation, participants were given a questionnaire to fill out that asked about the presence of any synaesthesia, previous use of drugs as well as visuo-spatial mental imagery. On the first day of experimentation, participants completed the sound-colour task at two timepoints, giving us individual measures for intra-subject consistency testing and an average colour for inter-subject consistency. The averaged colours for each day provided data for the sound-colour correspondences and LSD hallucinations. Participants always did the task in the non-LSD condition prior to the LSD condition, this ordering helped to familiarise participants with the task prior to LSD administration. The familiarisation is important as LSD can alter both the decision making abilities of users and their perception of incoming stimuli, so the process of familiarisation prior to LSD administration should reduce any ambiguities (Sinke et al., 2012). In addition, establishing the context of the sounds is important in predicting the colour correspondences likely to occur (experiment 1 and 2). On day 2 of testing, LSD was administered by intravenous injection, consisting of doses originally at 40mcg during pilot testing, however higher doses were administered on later participants (up to 80mcg) in order to enhance the probability of synaesthesia-like visual hallucinations in response to the sounds. Day 2 also had participants completing the task to completion twice.

5.10. Results

5.10.1. Presence of sound-colour hallucinations

In order to check that participant's undergoing the LSD condition did have visual hallucinations in response to the auditory stimuli, after each colour selection, participants rating how vivid each hallucination was on a four point scale from 1 (no experience) to 4 (absolutely clear image). A comparison of the average PAS scores revealed no significant differences in terms of their perceptual awareness scale scores, t (10) = -1.71, p = .122 with the non-LSD group having a mean score of 1.63 (SD = 0.50) and LSD group only marginally higher at 1.85 (SD = 0.75). Participants showed large variability in both control and LSD conditions, with some participants reporting consistently low (1.08) or high (2.35) visualisations for the non-LSD condition, and for some participants the LSD condition did not increase any reported hallucinations as measured by the PAS. Since participant's actively denied the presence of sound-colour synaesthesia, one explanation of this could be that some participants were reporting the vividness of their visual imagery, however while a correlation between visual imagery and PAS score is marginally significant for the LSD condition, $r = .627$, $p = .052$, this does not appear to be the case for the control condition, $r = .482$, $p = .158$.

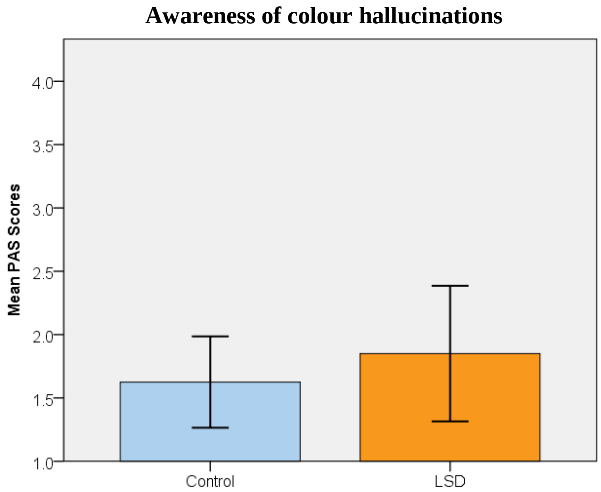


Figure 5.14. Mean perceptual awareness scale scores for colours chosen in response to auditory stimuli for participants on day 1 (control) and day 2 (LSD). Error bars indicate 95% confidence intervals.

5.10.2. Sound-colour correspondences

The average colour for each auditory stimulus within each condition (control, LSD) was used to establish the colour for analysis. Each colour provided four dependent variables, consisting of luminance, saturation, U* and V* scores.

5.10.2.1. Pure tone frequency

Increasing the frequency of pure tones increased the luminance of colours chosen, $F(5,45) = 2.91$, $p = .023$ (see figure 5.15), in line with prior studies (Marks, 1975). Contrary to experiment 2, increases in pitch did not have a significant effect on saturation, $F(5, 45) = 1.38$, $p = .249$. Also in contradiction to experiment 2, no significant main effects was found for increasing frequency on U* or V* dimensions, with $F(5, 45) = 0.31$, $p = .902$ and $F(5, 45) = 1.86$, $p = .121$, respectively. Indicating that prior pitch-saturation, pitch-U* and pitch-V* mappings may be masked when luminance is not controlled for. Finally, exploratory analyses for all colour dimensions yielded no significant main effects for LSD or interactions between LSD and frequency.

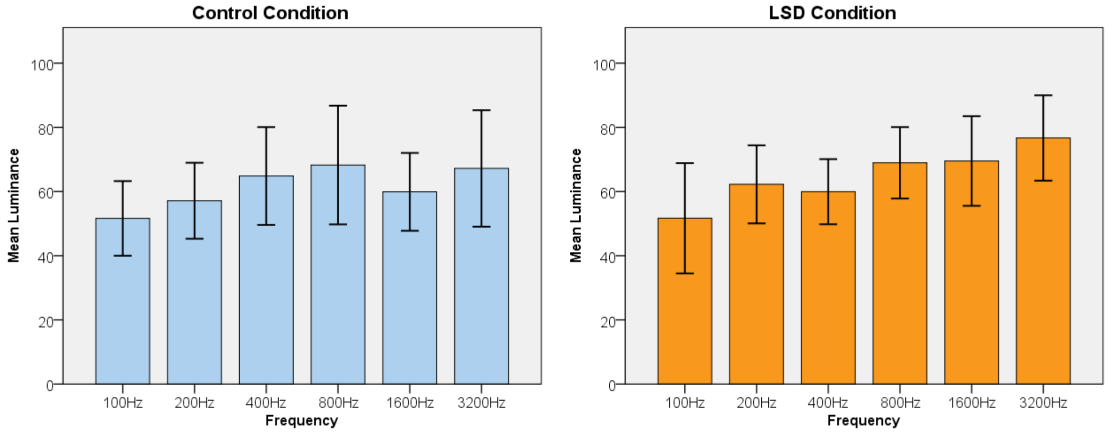


Figure 5.15. Mean luminance of colours selected for both the control (left graph) and LSD (right graph) conditions in response to different pure tone frequencies. Error bars indicate 95% confidence intervals.

5.10.2.2. Harmonics

A repeated measures ANOVA revealed a main effect of auditory stimulus on luminance, *F* (5, 45) = 2.99, *p* = .021, with increased disharmonics (noise) associated with increased luminance (see fig. 5.16) similar to previous findings by Lewkowicz & Turkewitz (1980). A main effect of increased harmonics increasing saturation in colours was also found, *F* (5, 45) = 9.33, *p* < .001, similar to experiment 1 (see fig. 5.17). No main effects of LSD or interaction effects were found for either the disharmonics-luminance or harmonics-saturation correlations. Exploratory analyses also found additional correlations between harmonics and V* dimensions, *F* (5, 45) = 2.98, *p* = .021, where disharmonic stimuli were deemed bluer while harmonic stimuli were yellower. No significant main effects were found for U* dimensions, in addition there were no main effects of LSD or interaction effects any for the U* and V* dimensions.

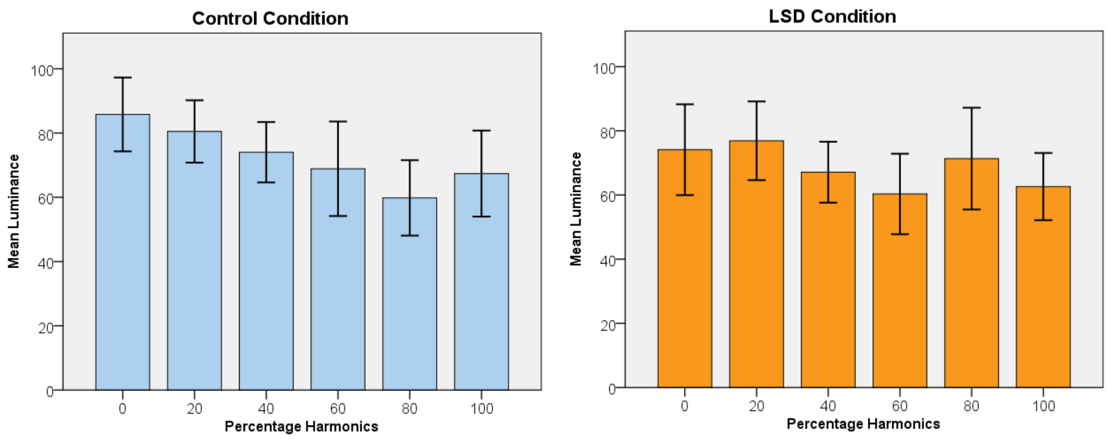


Figure 5.16. Mean luminance of colours selected for both the control (left graph) and LSD (right graph) conditions in response to different harmonic and disharmonic sounds. Error bars indicate 95% confidence intervals.

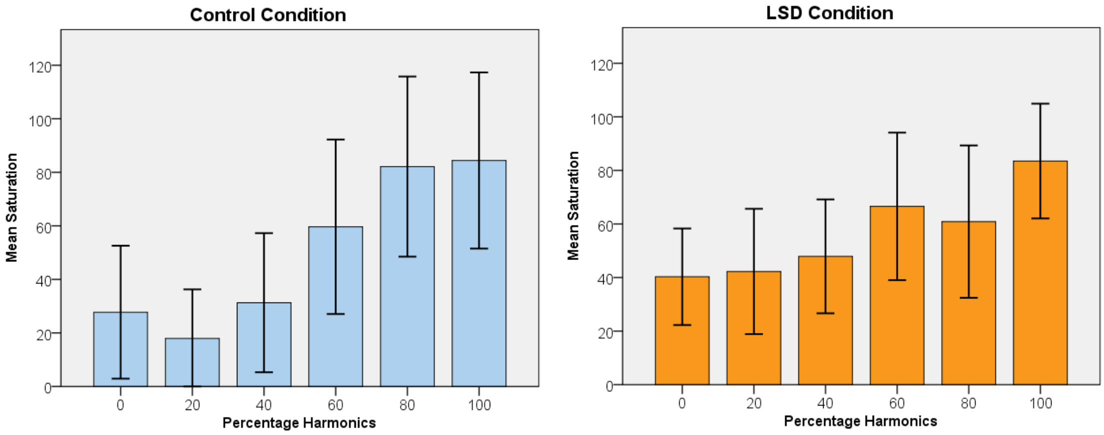


Figure 5.17. Mean saturation of colours selected for both the control (left graph) and LSD (right graph) conditions in response to different harmonic and disharmonic sounds. Error bars indicate 95% confidence intervals.

5.10.2.3. Vowels

Multi-level modelling for vowel-colour correlations allows an examination of whether specific colour dimensions (luminance, saturation, U* and V*) vary in response to either auditory brightness or compactness. Across the control and LSD conditions, auditory brightness was found to be significantly correlated with luminance, *t* (89.60) = 5.85, *p* = .018, with increased 'auditory brightness' resulting in more luminant colours. This is similar to previous observation that the presence of higher frequencies tend to lead to more luminant colour selections in rich timbral sounds (Fernay, Reby & Ward, 2012; Ward et al., 2006). Exploratory analyses also found another interesting correlation, with increased auditory brightness increasing the saturation of colours in both conditions to a marginally significant degree, *t* (98.78) = 3.84, *p* = .053, similar to the 'auditory brightness' to saturation mapping found for vowels (see section 5.7.8.3). No correlations between auditory brightness and U* or V* dimensions were found from exploratory analyses.

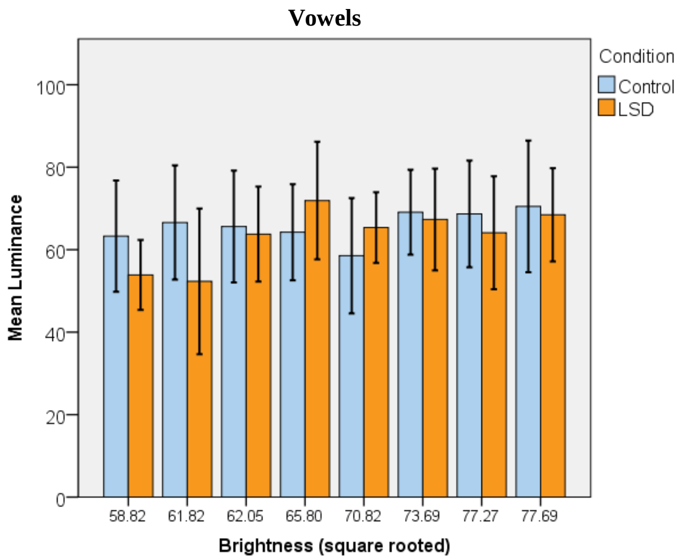


Figure 5.18. Luminance of colours selected for vowel sounds either in the control condition or LSD condition. Error bars indicate 95% confidence intervals. Error bars indicate 95% confidence intervals.

The compactness of the first two formants for vowel sounds has been found to predict colour correlations by placing vowel sounds in perceptually meaningful order (Marks, 1987; Moos et al., 2014; experiment 1 & 2). Multi-level modelling found marginally significant correlations between auditory compactness and the U* colour dimension, t (53.74) = -1.92, *p* = .061, where more compact stimuli yielded redder colours while more diffuse vowels yielded greener colours, similar to previous findings (experiment 1 & 2). No significant effects were found between auditory compactness and luminance, saturation or V* colour scores.

5.10.3. Consistency

5.10.3.1. Within participants

During each condition, participants went through the sound-colour matching task twice, one question is whether participants become more or less internally consistent with their own colour selections. The difference in CIE LUV scores between time 1 and time 2 were taken for the non-LSD and LSD conditions for comparison. A repeated measures t-test found that there was no significant difference for internal consistency while participants were either on or not on LSD, *t* (9) = -1.15, *p* = .280. As a result, LSD did not appear to create more consistent sound-colour pairings than when no LSD was taken.

5.10.3.2. Between participants

One possibility is that under LSD conditions participants may become more consistent with other members of the group if participants more readily utilise correspondences common to the group while disinhibited on LSD. An analysis of the cumulative distance between an individuals' CIE LUV scores for each auditory stimulus in comparison to other group members was conducted (lower scores indicate greater consistency). A repeated measures t-test found that there was a significant effect of LSD on consistency between group members, *t* (9) = -2.49, *p* = .035, however participants were more consistent with each other in the control condition (Mean = 8991.04, SD = 646.87) rather than the LSD condition (Mean = 9676.94, SD = 700.74).

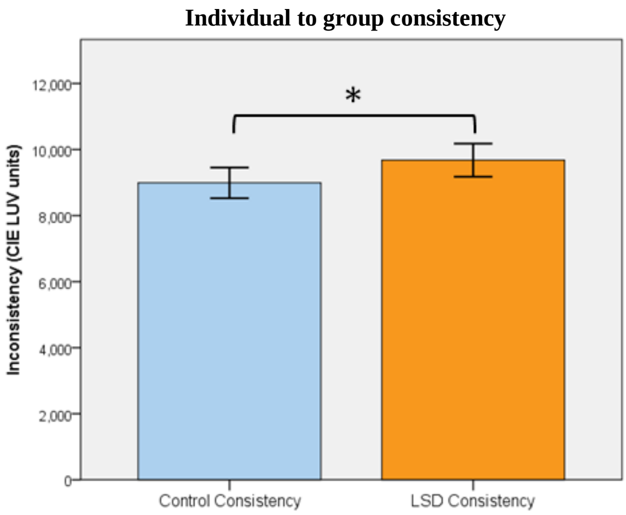


Figure 5.19. Individual to group consistency in CIE LUV units, lower scores indicate greater consistency. Error bars indicate 95% confidence intervals. Error bars indicate 95% confidence intervals.

5.10.4. Colour choices on LSD

The consumption of LSD can lead to a variety of perceptual disturbances on visual input. This may lead participants to select systematically different colours than when they are not on LSD. To investigate this possibility the average luminance and saturation were taken for each condition for comparison. For luminance, no significant difference was found, *t* (9) = 0.68, *p* = .516, likewise saturation also showed no significant difference with *t* (9) = 0.57, *p* = .584, from this we can conclude that colour selections were not systematically altered by LSD.

5.11. Experiment 3 summary

This report details the first systematic experiment on sound-colour correspondences while under the influence of LSD. The extent to which visual photisms were produced was assessed using standardised measures of conscious awareness, while colour selections were analysed for internal / external consistency, shifts towards particular colours and for previously observed correspondences. Despite moderate doses of LSD (40-80mcg) participants did not appear to report more visual photisms while under the influence of LSD (Luke & Terhune, 2013). Although a tendency in this direction is found, large individual differences make conclusions about any LSD photisms themselves tentative. Originally the experiment was devised to assess the qualitative attributes of LSD-induced hallucinations in response to auditory inducers, however measures assessing qualitative changes in conscious experience do not support the notion that LSD hallucinations occurred (Sandberg et al., 2009). Therefore any conclusions from the data must reflect on the nature of correspondences while under the influence of LSD instead. Beyond whether colour selections actually reflected LSD hallucinations, we investigated any systematic changes in colours chosen. Participants also did not choose more saturated or luminant colours on LSD, so LSD did not systematically lead participants towards certain characteristics of colour and any visual disturbances are unlikely to have a large effect here (Hartman & Hollister, 1963). Colour selections were also equally consistent within individuals irrespective of LSD consumption. Interestingly, while internal consistency remained stable, the colour selections between participants became further apart, indicating that colour selections were becoming more idiosyncratic on LSD. One potential explanation of this is that the ability to accurately discriminate different colours may have reduced on LSD inadvertently leading to more inconsistency (Abraham, 1982; Abraham & Wolf, 1988; Hartman & Hollister, 1963). In both the control and LSD conditions participants demonstrated previously observed correspondences for pitch-luminance, noise-luminance and harmonics-saturation. For vowel sounds, previously observed ‘auditory brightness’ changes lead to increased luminance and saturation of colours selected and the compactness of the first two formants lead to redder colour selections to a marginally significant degree. These findings indicate that correspondences are maintained under LSD, however since reports of actual colour photisms are not significantly different under LSD for our participants, the data cannot comment on qualitative changes in LSD hallucinations.

5.11.1. Correspondence mechanisms under LSD

Previous investigations of drug-induced synaesthesia utilising mild doses of LSD (1mcg/kg), found that auditory tones could elicit visual disturbances in just fewer than 50% of participants including increased luminance, geometric shapes and chromatic photisms (Hartman & Hollister, 1963). Luke and Terhune (2013) note additional informal experiments, case studies and surveys made in the literature that support this notion of induced audio-visual synaesthesia from LSD.

However, the lack of direct experiments makes comments on the qualitative attributes of these hallucinations difficult to determine. LSD also has many effects from both from a physiological and neural standpoint, serotonin agonists have been implicated in multiple forms of drug-induced 'synaesthesia' so the serotonergic system may play a strong role, likewise increases in cortical excitability and activity for the frontal lobe, limbic and paralimbic regions have also been found (Luke & Terhune, 2013; Sinke et al., 2012). Since LSD appears to have such wide reaching effects and qualitative changes in audio-visual perception, it becomes reasonable to both test whether the mechanisms that give rise to audio-visual correspondences remain intact and whether they influence the nature of these disturbances (Spence, 2011). Correspondences appear to use several mechanisms such as matching for intensity or emotional valence, which may involve regions affected by LSD such as the limbic system. Our experiment is able to provide evidence in favour of correspondence-mechanisms remaining stable, but cannot comment on whether these are utilised to form LSD hallucinations.

5.12. Overall discussion

Evidence is presented for a variety of audio-visual correspondences, this research reduces the substantial confound of varying luminance when looking to examine hue and saturation correspondences by utilising physically equiluminant conditions (Spence, 2011). However it should be noted that subjective brightness of stimuli are both dependent on the hue and saturation of colours observed, as well as subject to individual variation between participants as indicated in the Helmholtz-Kohlrausch effect (Nayatani, 1999). Further work may be considered on equibright colours that may instead vary in physical luminance to complement this data. That considered, further examinations of these correspondences reveal that some are rigidly linked to specific frequencies (i.e. yellow and the presence of frequencies above 800Hz), whereas others are flexible, either re-applying themselves to new contexts (i.e. pitch-saturation). It is suggested that some of these newly discovered correspondences are normally over-ridden by stronger pitch-luminance correspondences when they are not under equiluminant conditions. The mechanisms underlying correspondences are also shown to operate when under the influence of LSD, a hallucinogen that alters audio-visual perception and integration. This final discussion will examine some problems with prior research and how addressing these concerns have revealed a wide variety of correspondences, their tendencies and potential explanations.

5.12.1. Previous research

Previous studies attempting to examine audio-visual correspondences have included several flaws that make specific mappings difficult to view in isolation. The use of auditory stimuli have typically not controlled for interacting dimensions in pure tone stimuli such as pitch and loudness (ISO, 2003), or had a systematic way of classifying more complex timbral stimuli into a meaningful order (Ward et al., 2006), whereas for the visual correspondence, while the attributes of luminance and saturation have been covered in multiple papers (Marks, 1987; Martino & Marks, 1999; Melara, 1989; Orlandatou, 2012; Ward et al., 2006), saturation towards particular hues has only recently been analysed through the use of colour spaces based on human colour perception (Moos et al., 2014). Similarly, there have not been any previous experiments on hue or saturation while controlling for the effects of physical luminance (Spence, 2011). One potential criticism to consider is that subjective measures of visual brightness from observers could vary as a result of altering saturation or hue components such as those observed in the Helmholtz-Kohlrausch effect (Nayatani, 1999). While this effect may produce higher subjective brightness ratings for intense saturations of focal hues, maximum saturation was limited in this experiment so this effect may be minimised. Future studies could seek to address these concerns could either establish colours used had equal brightness through flicker fusion tasks, or alternatively modify the colour space that is manipulated to account for general tendencies such as through using Ware-Cowan equations or psychophysical data (Fairchild, 1998, p. 142; Pridmore, 2007). The effect of saturation on the perceived brightness of equiluminant colours also varies as a function of hue. For example, saturated blues show an increased brightness relative to equally saturated red hues. These effects are also subject to large individual variation (Ayama & Ikeda, 1998). Despite these reservations our paper helps to address these concerns while further examining the flexibility of these correspondences to different auditory contexts as well as breaking down the characteristics of complex sound-colour correspondences.

5.12.2. Observed sound-colour correspondences

Pure tone sounds were found to exhibit positive frequency-saturation mappings across both 100-3200Hz and 440-880Hz frequency ranges. This demonstrates not only a new correspondence likely obscured in non-equiluminant conditions, but that it is also flexible in applying itself to different contexts. Similarly, higher frequency ranges in bandpassed vocal sounds also demonstrated frequency-saturation mappings showing that this is also applicable to more complex timbral sounds. Besides frequency, the loudness of pure tones was also related to higher saturations. For sounds that shared the same frequency range, complex sine wave sounds had the loudness of the higher frequencies ('auditory brightness') correlated to increases in colour saturation. This was also found for vowel sounds across speakers with higher frequency formants also increasing saturation, this correspondence was even maintained under the influence of LSD. Also as seen in experiment 1, increases in colour saturation are also driven by harmonics. So from pure tones to vowel sounds, increases in the loudness of relatively 'high' frequencies correlate to increases in colour saturation, when luminance is controlled for.

Certain auditory dimensions corresponded to the saturation of specific hues. Higher frequencies correlated to yellow-saturations for pure tones, complex sine wave sounds and bandpassed vocal sounds for frequencies above 800Hz. Furthermore, complex sounds with contained frequencies of 800Hz and above had participants consistently selecting yellow-hues. These findings in equiluminant conditions suggest that previous experiments finding correlations between frequency and yellow hues (Orlandatou, 2012; Simpson et al., 1956) are not the sole result of frequency-luminance correspondences (Spence, 2011). Instead our research suggests that there is a separate correspondence between yellow hues and frequencies above 800Hz. Our own research into vowels demonstrated a correlation between the compactness of the first two formants and red-saturation, as well as a new correspondence between auditory-dullness and red-saturation. Despite predictions by Marks (1974) that auditory brightness should correspond to yellow hues, our data did not support this prediction in experiment 1. Similarly, Moos et al. (2014) found that neither F1 nor F2 correlated with yellow-saturation either. Interestingly our exploratory analysis on vowels across speakers found a marginally significant effect of auditory brightness on yellow-saturation. These differences might be due to the inclusion of higher frequency formants in female-typical voices in experiment 2. It was found by comparing vowels and complex-sine wave sounds derived from their formant positions did not correspond to the same colour correlation. This indicates that red-green correlations may be the result of the more complex timbre present in vowel sounds.

5.12.3. Categorising correspondences

Previous theories regarding the explanation of correspondences have principally examined the difference between lower-level automatic correspondences and higher-level decisional correspondences (Spence, 2011). However as the current three studies did not involve speeded classification tasks, but rather freeform matching, it is difficult to comment on the level of automaticity for any audio-visual pairings observed here. Instead two distinct patterns of correspondences appeared to emerge from our testing, correspondences that are sensitive to changes in context and those that are not, which we have referred to as flexible and rigid respectively. We observed flexible correspondences for pitch-saturation and compactness-red correspondences, with correlations between the two adapting to fit the range set by the stimuli. In contrast to this, rigid patterns emerged for high frequencies (above 800Hz) and yellow hues across a wide variety of conditions.

5.12.3.1. Flexible correspondences

The different characteristics of flexible and rigid correspondences may indicate different underlying mechanisms. For flexible correspondences, discrete values for the auditory stimulus might be abstracted into simpler representations of magnitude, where a stimulus is rated as relatively low to high based on where the upper and lower bounds of stimulation are in a given context. One such example of this are the pitch-saturation mappings found to re-apply themselves to different frequency ranges. Walsh (2003) proposes such a mechanism to abstract magnitudes between seemingly independent qualities (time, space & quantity) in the parietal cortex. Of interest to the present research is that disruption of the intraparietal cortex can eliminate cross-modal integration (Bien et al., 2012), if flexible correspondences are based on magnitude matching in the parietal lobe, then it would be predicted that these correspondences would be reduced through parietal disruption similar to disruptions seen to developmental synaesthesia (Esterman et al., 2006; Muggleton et al., 2007). Flexible correspondences also seem to be more influenced by the presence of competing visual dimensions, pitch-saturation is a correspondence that appears strongly in equiluminant conditions (experiment 2), but only few studies have demonstrated this previously, either through rarely observed linear relationships (Orlandatou, 2012) or by non-linear mappings that reduce as luminance is increased (Ward et al., 2006). Instead, pitch-luminance correspondences appear to dominate pitch-saturation mappings, either this is a favoured visual dimension for magnitude matching, or pitch-luminance is a rigid correspondence less affected by context (Thornley Head, 2006; Ward et al., 2006) and so takes precedence over correspondences that need to be abstracted into magnitudes first.

5.12.3.2. Rigid correspondences

Rigid audio-visual correspondences appear to be less influenced by their context, with specific auditory characteristics linked to a single visual dimension. For a variety of stimuli, we repeatedly noticed sounds that had frequencies over 800Hz being paired with yellow hues, this was true for pure tones, complex sine wave stimuli and vocal sounds (both varying in frequency and brightness). In addition, increased loudness for higher frequencies in vowel sounds also demonstrated this correspondence. The mechanisms behind this are unclear at present, they may be based on structural similarities in cortical representations between pitch and yellow, or through learned statistical correlations with the environment (Spence & Deroy, 2012), for which pitch-yellow relationships may be related to pitch-luminance associations despite our equiluminant stimulus presentation. A psychological factor that should also be noted is that while equiluminant stimuli were presented, there may be tendencies for participants to associate presented colours with the purest exemplars of that given colour (i.e. a dull yellow may remind them of a prototypical bright yellow) which may provide a route for pitch-luminance associations to manifest. Counter to this is that the opposite might be expected to occur, since a prototypical blue is darker than yellow, the blues presented might be more luminant than their prototypical exemplar, which might encourage its choice to express pitch-luminance associations. Another potential explanation is that of emotionally matched stimuli, with higher pitches and yellower hues linked by positive emotional valence (Palmer et al., 2013; Sebba, 1991), if this were the case we would expect pitches above 800Hz to be deemed emotionally positive.

5.12.4. Correspondences and the wider context

The relationship between sound and colour has inspired discussion examining their structures in physics (Newton, 1706), their aesthetic appeal in art (Jewanski, 2010) and their processing in psychology (Spence, 2011). Through understanding what auditory characteristics drive correspondences to specific visual features and their flexibility, these factors help us to understand the mechanisms that could explain these mappings. Furthermore these intuitive mappings can also pave the way in designing optimal solutions for communicating visual information using sound. The use of equiluminant conditions in particular have illustrated previously hidden correspondences that may have been masked during previous investigations as well as disentangled the influence of luminance in explanations of these correspondences.

6. Sound-Hue Correspondences Reduce Colour-Based Errors in Sensory Substitution

6.1. Abstract

Visual sensory substitution devices (SSDs) can represent visual characteristics through distinct patterns of sound, allowing a visually impaired user to access visual information. Previous SSDs have tried a wide variety of visual to auditory mappings, however to date there has not been a systematic attempt to determine if there are optimal cross-sensory pairings for the end user. One potential answer may be the use of cross-modal correspondences. These are widespread intuitive pairings between specific sensory features in different modalities. Since congruent pairings are processed faster, this may make correspondences useful in reducing the difficulty of learning to use SSDs. Recent evidence has also established correspondences between specific auditory characteristics and saturation towards specific hues. To test the utility of correspondences, a tablet device was created that converts a single point of colour into variations of sound, we examined the colour-sound associative memory and object recognition abilities of users who had their device either coded in line with or opposite to sound-hue correspondences. Improved colour memory and reduced colour-errors were made by users who had the correspondence-mappings. Interestingly these users also made fewer luminance errors, a mapping that was identical across groups. Furthermore, the colour-sound mappings that provided the highest improvements during the associative memory task also saw the greatest gains for recognising realistic objects that also featured these colours, indicating a transfer of abilities from memory to recognition. This indicates that intuitive sound-hue mappings make colour identification easier as well as allow users to explore other aspects of an image more thoroughly to reduce other errors. These findings are discussed with relevance for both colour and correspondences for sensory substitution use.

6.2. Introduction

Blindness and other forms of visual impairment affect approximately 285 million individuals across the world, for which 20% cannot be avoided or cured through contemporary medical interventions (WHO, 2014). As the vast majority of visual deprivation occurs during disruption to the normal visual pathways, one approach is to bypass the visual pathways and deliver visual information to the remaining modalities for the brain to process. This can be achieved through the use of 'sensory substitution devices' (SSDs) that systematically translate visual information (e.g. spatial position, luminance and colour) into dimensions of sound (e.g. variations of pitch, loudness and timbre). This process allows SSD users to decode these sounds back into a coherent mental visual picture for use as a visual aid in recognising, exploring and interacting with the 'visual world.' However despite the apparent utility of such techniques to extend a users' understanding of the visual world, these devices are rarely used outside of experimental settings. Part of this is due to their lack of availability, such as with a lack of knowing about these devices, concerns about their appearance in public, their cost, difficulty in setting up, lack of systematic training and over-riding of important auditory cues in daily life (Maidenbaum, Abboud & Amedi, 2014). Other concerns are for the utility of the devices themselves, such as the difficulty of solving visual tasks through the use of another modality. This difficulty in SSD users bridging this 'cross-modal gap' can be alleviated in several ways, first there is the reduction in the amount visual information, then there is selecting the most distinctive auditory dimensions for translation and finally there is the use of intuitive cross-modal correspondences (Hamilton-Fletcher & Ward, 2013; Chapter five).

6.2.1. Colour spaces

Visual dimensions for use in sensory substitution typically involve the vertical axis, horizontal axis and luminance, this reduction in visual information simplifies the translation and re-interpretation of the sounds. However, recently dimensions of colour have been incorporated into sensory substitution (Hamilton-Fletcher & Ward, 2013). Colour itself aids in scene segmentation and object identification at the low visual resolutions typically used in SSDs (Brown, Simpson & Proulx, 2014; Torralba, 2009), not only facilitating behavioural tasks involving colour (Abboud, Hanassy, Levy-Tzedek, Maidenbaum & Amedi, 2014) but allowing access to a shared lexicon of colour with the sighted. Since encoding the dimensions of colour take up valuable auditory dimensions that could be reserved for representing spatial dimensions, several reductions in colour-complexity are typically taken, including chunking colours together as categories, using simpler models of colour (e.g. RGB / HSL) or segmenting an image prior to conversion (Bologna, Deville, Pun & Vickenbosch, 2007). In comparison, some devices reduce the spatial complexity to reserve auditory dimensions for colour dimensions, such as the See ColOr which does not encode the vertical axis or the Kromophone which removes both the horizontal and vertical axes (Bologna et al., 2007; Capalbo & Glenney, 2009). Some devices that use all auditory dimensions for representing colour are the Soundview and Creole SSDs (Doel, 2003; Chapter four), in both of these devices, a single point of colour is translated into sound, with the spatial axes chosen via finger movements on a tablet device. Despite this focus on colour representation, neither of these devices has previously used a human perception-based model of colour perception such as CIE LUV colour space. These representations feature several advantages over alternative colour spaces through representing accurate perceptual distances, focal/non-focal colours and colour opponency. The closest example in SSDs is that of the ColEnViSon, which uses an alternative description of CIE LUV space; however the complexities of this are reduced through the use of clustering colours around common descriptive terms (Ancuti, Ancuti & Bekaert, 2009). So far, there are no devices that allow an examination of the complexities of colour in line with human colour perception, recent revisions to the Creole hope to rectify this in order to examine specific auditory representations of colour.

6.2.2. Sound-colour mappings

There are two main strategies for choosing optimal sound-colour pairings for use in sensory substitution devices. The first characteristic is that of auditory distinctiveness, whereby the characteristics that make a given sound distinctive are different from that of other sounds used, the effect being that one sound is not mistaken for another (Grey 1975; Wessel, 1979) and that these distinctive sounds can be used to represent distinct colour categories (Maidenbaum et al., 2014). While distinctive auditory dimensions may make the auditory classification easier, they do not necessarily help in the identification of the visual property it represents without explicit training on the mappings used. The second factor is that of cross-modal correspondences, whereby certain visual and auditory dimensions are treated as more 'equivalent' by the brain than their opposite pairings. For instance, higher auditory pitches are matched with higher visual luminances (Ward, Huckstep & Tsakanikos, 2006), and that processing of these stimuli is faster when these are congruently rather than incongruently matched (Marks, 1987; Martino & Marks, 1999; Melara, 1989). These mappings may be innate, based on statistical regularities in the environment, language, or shared processes such as emotional valance or arrived at through explicit decision making (Palmer, Schloss, Xu & Prado-León, 2013; Spence, 2011). The use of cross-modal correspondences have been utilised for some common SSDs such as the vOICe (Meijer, 1992), using pitch-height for the vertical axis, interaural differences for the horizontal axis and loudness for luminance (Collignon, Lassonde, Lepore, Bastien & Veraart, 2009; Lewkowicz & Turkewitz, 1980; Walker et al., 2010). The use of correspondences for colour representations have been more muted, with several devices using loudness-luminance correspondences (Capalbo & Glenney, 2009; Maidenbaum et al., 2014), pitch-saturation (Bologna et al., 2007) and loudness-saturation (Capalbo & Glenney, 2009; Else, 2012). However representations of specific hues have been more arbitrary, typically involving variations in pitch or timbre typically with reference to auditory distinctiveness but not correspondences (Hamilton-Fletcher & Ward, 2013). Some auditory characteristics have been found to be associated with hue and saturation, but have not yet been utilised. For specific hues, certain vowel sounds are more commonly associated with red or green (Marks, 1975; Miyahara, Kode, Sekiguchi & Amemiya, 2012; Moos, Smith, Miller & Simmons, 2014; Wrembel, 2009; Chapter five), while high and low complex sinewave tones are associated with yellow and blue respectively (Orlandatou, 2012; Chapter five). For saturation, the presence of harmonics are associated with more saturated colours (Ward et al., 2006; Chapter five) and noise with desaturation (Lewkowicz & Turkewitz, 1980; Chapter five). These sounds provide a way of examining how the use of correspondences may affect the use of colour in sensory substitution devices, while keeping the auditory distinctiveness of the mappings consistent. As congruent audiovisual correspondences are associated with faster processing and higher integration (Marks, 1987; Martino & Marks, 1999; Melara, 1989), the use of these would be expected to come with a variety of behavioural benefits. These may include faster memorisation of the mappings as well as fewer colour mistakes and less cognitive demand for solving a visual task utilising colour.

6.2.3. Hypotheses

In order to test the hypothesis that specific hues and saturations are better represented using correspondences rather than arbitrary mappings, a tablet SSD that utilises CIE LUV colour space was created. This colour space allows the application of colour opponency, so the 'opposite' of red would be green on the CIE U* dimension. The auditory representation for a given colour would be either in line with, or opposite to previous experiments. The expectation is that the 'normal' mapping would increase colour memory, decrease colour-errors and as a result improve scores on visual identification tasks over users with a 'reverse' mapping.

6.3. Method

6.3.1. Participants

Forty-two participants were recruited using 'Sona' the online recruitment tool for the University of Sussex. Participants who indicated the presence of any colour blindness, uncorrected hearing / vision impairment or sound-colour synaesthesia were excluded from participation. Across all groups, the participants age ranged between 18 to 24 years old (Mean = 19.40, SD = 1.42). Thirty-two of the participants were female, eight participants were left handed, and all participants were undertaking education at an undergraduate level bar one who was completing education at postgraduate level. Participants were either paid for their time or received course credits. Ethics were approved by the Life Sciences and Psychology Cluster-based Research Ethics Committee at the University of Sussex. Due to technical issues one participant was only able to complete the colour memory task and not the object identification task.

6.3.2. Design

The experiment used an independent measures design, with participants either having their Creole SSD programmed with previously observed sound-hue/saturation correspondences (henceforth referred to as the 'normal' group) or the opposite of these sound-hue/saturation correspondences (referred to as the 'reverse' group). For the colour memory task the dependent variable was the percentage of colours correctly identified. For the object identification task, dependent variables were the amount of correctly identified objects (in this case, fruits and vegetables). Further examination of the object task responses gives us the types of errors participants made, with the amount of colour errors (e.g. mistaking red for green) and luminance errors (e.g. mistaking bright for dark lighting). The task score data was analysed using independent T-tests for the colour memory task and object identification task. A set of T-tests was used to analyse both of the types of object identification errors, with an examination of both 'colour-based' errors and 'luminance-based' errors across groups.

6.3.3. Materials

6.3.3.1. The Creole Sensory Substitution Device

The Creole is a tablet device that allows users to explore a normally 'visual' image through the tactile exploration of the 'image' on the device's surface. The user cannot see any image displayed on the device so the only way to understand an image is through the tactile exploration's auditory feedback. The Creole works by turning a single colour underneath the users' fingertip into variations of sound for the user to listen to. The variations in sound are informative as to the colour underneath the fingertip and so by exploring the image over time the user can chart the variations of colour in multiple spatial locations in a serial fashion to understand what the image might represent.

The 'Creole' device combines a client program which is run on a pantech 4G LTE tablet on an Android OS which identifies the X and Y position of the user's fingertip. This information is sent via a USB connection to a Lenovo T520 laptop that runs the main application. The X and Y positions by the user selects the pixel of the target image to be sonified based on the algorithm described below. In our procedure, only one image is presented during each application use, so after an image is used, the application is shutdown and a new image is started with the application. This process also provides a 'starting tone' for the user to let them know a new image has been loaded on the device without any clues as to what the image is. The online processing of audio and tactile output from finger explorations over the image was done using an Intel HD graphics 3000 GT2+ graphics card. Audio stimulation was generated from the laptop program using a Conexant 20672 SmartAudio HD soundcard outputted through commercial earbuds. The program for the tablet and laptop were coded in C# using Microsoft Visual Studio 2012.

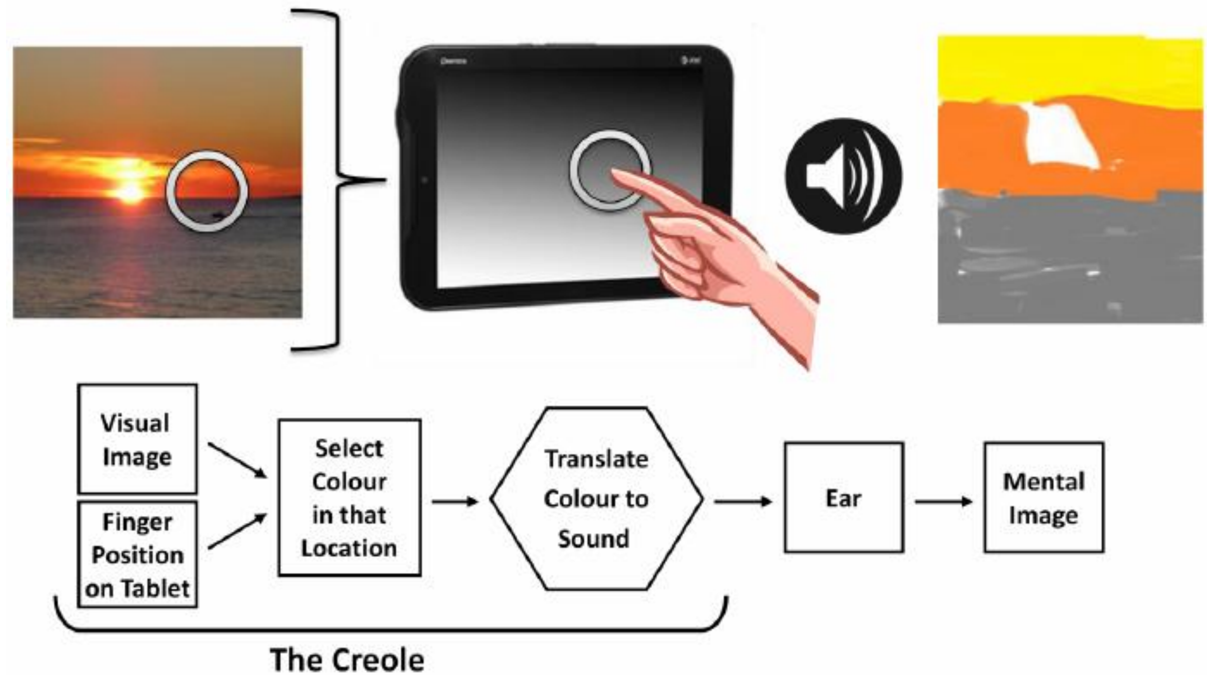


Figure 6.1. The Creole. Visual images are stored on the device but not shown to users. Touching the device relays co-ordinates of the finger to its equivalent position on the image and selects the pixel at that location. The pixel's colour values are then translated into patterns of touch and sound for the user. Through explorations over time, the user can build up a mental image of the distribution of colours in the image.

6.3.3.1.1. Creole colour space

Unlike most previous colour sensory substitution devices, the Creole uses a form of colour-space that is constructed to replicate veridical colour perception in a colour space known as CIE LUV. Each pixel's RGB colour values (representing red-saturation, green-saturation and blue-saturation) are translated into CIE LUV colour space which contains L* (white to black), U* (red to green) and V* (yellow to blue) dimensions as well as CIE LCH, from which a chroma (or saturation) score is derived (Tkalcic & Tasic, 2003). Each of the CIE LUV dimensions and the chroma dimension of LCH space are transformed into proportions of: whiteness, greyness, blackness, red-saturation, green-saturation, yellow-saturation and blue-saturation. For proportions representing saturation, a maximum upper limit was determined from the highest amount of red, green, blue and yellow saturation across all of the fruit/vegetable pictures, proportions under 0.20 were treated as zero. This allowed minor saturations to be ignored in favour of representing desaturation, as well as for participants to experience the whole proportion range when exploring colours in a natural image. The advantage of exploring the whole proportion range is that for novices using the device the variations in auditory output are greater than one which treats the maximum U* or V* values as the highest proportion points for which users are unable to experience these colour representations in the pictures of natural stimuli. The transformations that turn CIE LUV and LCH colour values into logarithmically scaled proportions of whiteness, greyness, blackness, redness, greenness, yellowness and blueness are listed below:

Ranges of values for natural stimuli used: L (0 to 100), C (0 to 180), U (-90 to +175), V (-135 to + 110).

$White_{proportion} = L^4 / 100000000$

$Black_{proportion} = (L – 100)^4 / 100000000$

$Grey_{proportion} = ((L * (L – 100) / 1000) / 2.5)^4 * ((180 – C) / 180)^4$

$Red_{proportion} = (U + 90)^4 / 493155625$

$Green_{proportion} = (U – 175)^4 / 493155625$

$Yellow_{proportion} = (V + 135)^4 / 3603000625$

$Blue_{proportion} = (V – 110)^4 / 3603000625$

If any proportions exceed 1 then they equal 1, if they fall under .20 they equal 0.

6.3.3.1.2. Creole sonification method

The sonification mappings used in the normal and reversed conditions are summarised in Table 6.1. For representing the presence of whiteness and blackness as separate entities, a high pitched (3520Hz) and low pitched tone (110Hz) represented high and low luminances respectively. This mapping was chosen based on previously observed pitch-luminance correspondences (Hubbard, 1996; Ludwig, Adachi & Matsuzawa, 2011; Marks, 1974; Ward et al., 2006). Previous studies have also found that particular sounds are associated with specific hues or saturations. Saturation and desaturation have been found to be related to harmonics and disharmonics in auditory stimuli respectively (Chapter five), so de-saturated colours are represented by bandpassed white noise (100 to 3200Hz). For the saturation of any hue, the sound representing the hue would need to be a harmonic correspondence. The presence of yellow was represented by three high-pitched frequencies making a C-major chord (1047, 1319 & 1568Hz) while blueness was represented by three low-pitched frequencies making a C-minor chord (262, 311 & 392Hz). These were chosen based on Chapter five's finding that bands of high-pitched frequencies were related to yellowness and low-pitched with blueness in equiluminant conditions (see also Orlandatou, 2012; Simpson, Quinn & Ausubel, 1956). Redness is represented by a male vocal vowel sound of /u/ while greenness is represented by a male vocal vowel sound of /i/. The vocal stimuli were the same as the stimuli that participants rated as the most red and green respectively in Chapter five (also see Marks, 1975;

Miyahara et al., 2012; Moos et al., 2014; Wrembel, 2009). Finally in order to communicate the increased presence of whiteness, blackness, greyness, yellowness, blueness, redness or greenness, the auditory dimension of loudness was used. Loudness is a dimension that has been both related to increased brightness as well as saturation (Giannakis, 2001; Lewkowicz & Turkewitz, 1980; Marks, 1974, 1987; Chapter five). Firstly, so that loudness could be easily compared across multiple sound stimuli, the maximum volume for each sound was loudness-equalised (so higher pitched tones had their amplitude reduced, so that high and low pitched tones sounded equally loud at their maximum values). This new maximum loudness value represents the highest proportion of colour-magnitude available to users.

Finally, for the ‘reverse’ condition, the sounds representing yellow and blue were swapped, red and green were swapped and finally grey was represented using harmonic stimuli (complex harmonic tones of 100, 200, 400, 800, 1600 and 3200Hz), while white and black sounds remained the same. Due to the substantial prior evidence that auditory pitch can affect visual luminance processing (Spence, 2011), the processing effects of switching which frequencies represented luminance was deemed uninteresting as well as potentially obscuring the less investigated potential effects of saturation or hue correspondences on colour processing. In the interest of exploring this area in isolation, only hue and saturation correspondences were candidates for reversal.

The four unique hues are primarily orientated with either the CIE U* or V* dimensions, each of these hues primarily consist of increased values along one of these axes, for example a focal yellow consists of a positive V* value while U* remains neutral. This results in a pure sound representing yellow to the user (1047, 1319 & 1568Hz tones). The addition of positive U* values to this deviates from a focal yellow by adding red (represented by /u/ vowel sounds), that in combination create orange (a combination of 1047, 1319 & 1568Hz tones with an /u/ vowel sound). As such, unique hues sound pure in that they only contain one sound in isolation, while non-unique hues such as orange have a sound that mixes the constituent hues (and sounds) together. This mimics veridical colour perception, where unique hues are seen as not containing any saturation of other hues, whereas non-unique hues are perceived as being a mixture (Miyahara, 2003; Regier, Kay & Cook, 2005).

Table 6.1. Sound-colour mappings for the 'normal' and 'reverse' correspondence conditions.

Colour	Normal Mappings	Reverse Mappings
White	3520Hz pure tone	3520Hz pure tone
Grey	**100**-**3200**Hz **noise**	**100, 200, 400, 800, 1600 and 3200**Hz tones
Black	110Hz pure tone	110Hz pure tone
Red	/**u**/ vowel sound	/**i**/ vowel sound
Green	/**i**/ vowel sound	/**u**/ vowel sound
Blue	**262**, **311** and **392**Hz tones	**1047**, **1319** and **1568**Hz tones
Yellow	**1047**, **1319** and **1568**Hz tones	**262**, **311** and **392**Hz tones

6.3.4. Procedure and tasks

After participants read through the information sheet, filled out the consent form and demographics form. Participants were introduced to the Creole through a PowerPoint presentation on a desktop 19 inch flat screen monitor running at 1024 by 1080 resolution with 60Hz refresh rate. The presentation informed users that the tablet device 'displays images' but the user will only be able to hear the images rather than see them. Users were informed that they can listen to the image by dabbing their finger on the tablet, which turns the colour underneath their finger into sound for them to listen to and that specific colours make specific sounds so it becomes possible to figure out what the colour underneath their finger is. Then by exploring an image over time, users can figure out what an image might be from location of the colours through sound. The explanation of the more technical aspects of the Creole showed users that the Creole understands colour as consisting of seven colours (white, grey, black, red, green, blue and yellow), and that each of these seven colours has a unique sound associated with them. Then this colour-sound mapping is revealed to users (either along 'normal' or 'reverse' mappings), with a verbal explanation describing the sounds and that they will be able to listen to examples shortly. Then the process of colour mixing from these seven basic colours is explained to users, with the loudness of the sounds indicating the increased presence of one of the seven basic colours. In addition to this, the presence of multiple sounds (and their relative volumes) indicate that these basic colours are being mixed together, such as with dark/bright saturated colours or a more complex non-unique hue, such as orange. After participants' questions are answered and are confident with the explanation, the colour training stimuli are presented for exploration by the user. These are presented non-visually on the Creole tablet (running from a Lenovo T520 laptop) for active exploration by the user. All images in all conditions are 1080 by 720 pixels in size, and presented in a serial fashion to the user.

6.3.4.1. Colour training

Figure 6.2. Creole colour training materials, left image allows users to explore the main seven colours, middle image allows users to explore bright and dark variants of the four unique hues, and right allows users to explore complicated colour combinations when two unique hues mix.

Colour training stimuli consisted of three images allowing an exploration of either the primary seven colours used in the Creole, the second image allowed an exploration of light and dark variants of the four unique hues and the final colour training image allowed users to explore mixes of the unique hues. These images were presented in a serial fashion to the user (see fig. 6.2) on both the participant's computer screen and through sound (minus the text and visual reference) on the Creole. Participants were able to freely explore these until they felt comfortable with the sounds, typically taking five minutes to go through these three training items.

6.3.4.2. Colour Memory Task

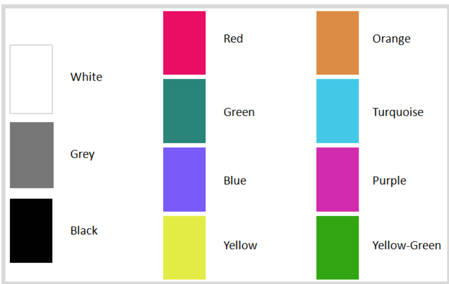


Figure 6.3. Visual image of potential individual colours presented on the Creole that the participant can select from.

The first task was the colour-sound associative memory task, where users were presented with one of eleven colours on the Creole and asked to select which of the onscreen colours (see fig. 6.3) was being sonified. Participants could take as long as they wanted across the thirty-three trials, with each colour being presented three times in a randomised order. Participants verbally identified which colour they believed they were listening to, and each answer received verbal feedback from the experimenter, with either 'yes, that is correct' or 'the sound you were listening to was <colour>'

with a quick description of what sound represents that colour. For the colour memory task, the visual presentation on the computer for the participant showed the eleven colours used in the task, consisting of the seven primary colours the Creole uses and four complex hues. Luminance for some of these hues varied, with yellow, turquoise, purple and 'yellow-green' having luminances above L = 50, so not all colours can be done from rote memorisation from training, as some inference is required. On the Creole itself a block of one of these eleven colours was presented for the participant to listen to, so that anywhere a participant touched the Creole elicited the same sound. Once the colour memory task was complete, shape training was given.

6.3.4.3. Shape training

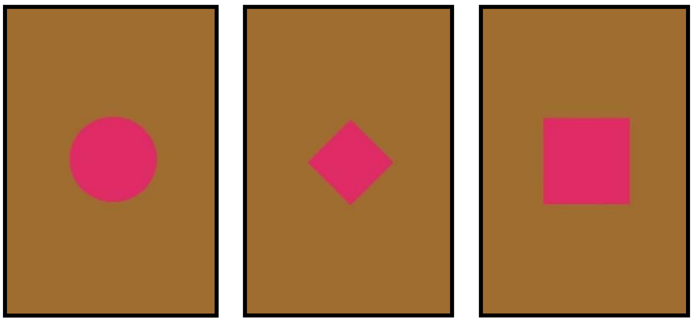


Figure 6.4. Visual image of three potential shapes that could be presented on the Creole for participants to explore and listen to.

Shape training was employed in order to have participants practice how combining information about colours and their locations can be used to identify shapes. Stimuli consisted of a red circle, diamond or square on a brown background. By plotting out which parts of an image belong to the 'target' and which belong to the 'background' through auditory stimulation, users can figure out which of the three objects is represented on the Creole. The visual presentation consisted of all three objects on the PC screen, while one of the three objects was loaded onto the Creole for participants to listen to, participants were asked to explore the image and pay attention to the 'target' colour, the 'background' colour and then to concentrate on areas in which the three objects differ. By exploring spatial locations in which the shapes vary from one another users could figure out which of the three shapes was being listened to on the device. Three trials of this were performed with feedback for training purposes. Participants were given feedback for each of their answers on the three stimuli and no time limit was imposed. Finally, participants were given the object identification task.

6.3.4.4. Object identification task

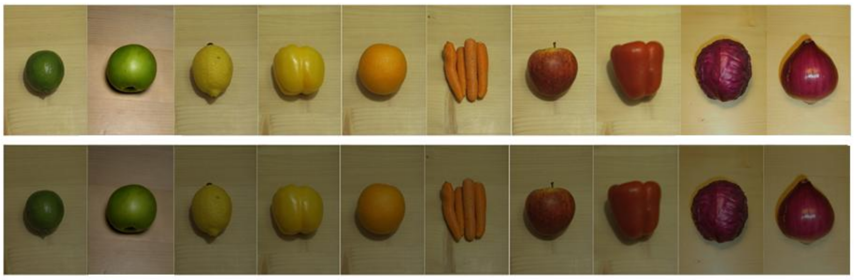


Figure 6.5. Visual presentation of object stimuli, consisting of twenty fruit/vegetable split into five different colours, and two different luminance conditions. Participants were presented with one of these stimuli on the Creole to listen to and determine which visual image they believed it represented.

In order to test the utility of the sound-colour mappings an object discrimination task was utilised, where a single image of a fruit or vegetable is presented on the Creole, while twenty different images of fruit and vegetables are presented on screen. Participants explore the image in order to make the determination of which of the twenty images they are listening to on the device. The twenty potential images were presented on screen as shown in fig. 6.5 consisting of two green, yellow, orange, red and purple fruit/vegetables, in either bright or dark lighting conditions. These stimuli provided ecologically valid colours, with variations of colour (e.g. darker coloured shadows) in line with what is expected in a natural environment. This was done to supplement tests on abstract representations of colour (e.g. computerised colours) to apply to a more realistic use of colour. The twenty trials consisted of a single randomly selected image represented on the Creole for participants to identify in a 20 alternative forced choice task. Participants were reminded that due to the random selection the same image would be able to be represented more than once. Participants were reminded to identify the fruit/vegetable by brightness of the image (light, dark distinction), and by the colour and shape of the fruit/vegetable. . Participants were given feedback on their answers, the responses that were given were either “yes, that's correct” for a correct answer, or “the correct answer was the <dark/light> <colour> <fruit>” (e.g. *dark red apple*). Due to the difficulty of the task, participants were encouraged to use the feedback for any incorrect answers to help them guide future searches as well as to concentrate on areas in which fruit/vegetables differ from one another to maintain focus and motivation for the task. For the first image presented participants were able to take up to three minutes before answering, and subsequent trials participants were asked for their answer after a minute. After the completion of this final task, participants were debriefed, thanked for their time and effort and either paid in course credits or £10.

6.4. Results

6.4.1. Colour memory task

The colour-sound memory task consisted of eleven different colours, each of which is presented three times to the user. This gives us both a cumulative score for how well each colour is remembered for each group (see fig. 6.6) and the overall amount of colour errors across all colour stimuli (see fig. 6.7).

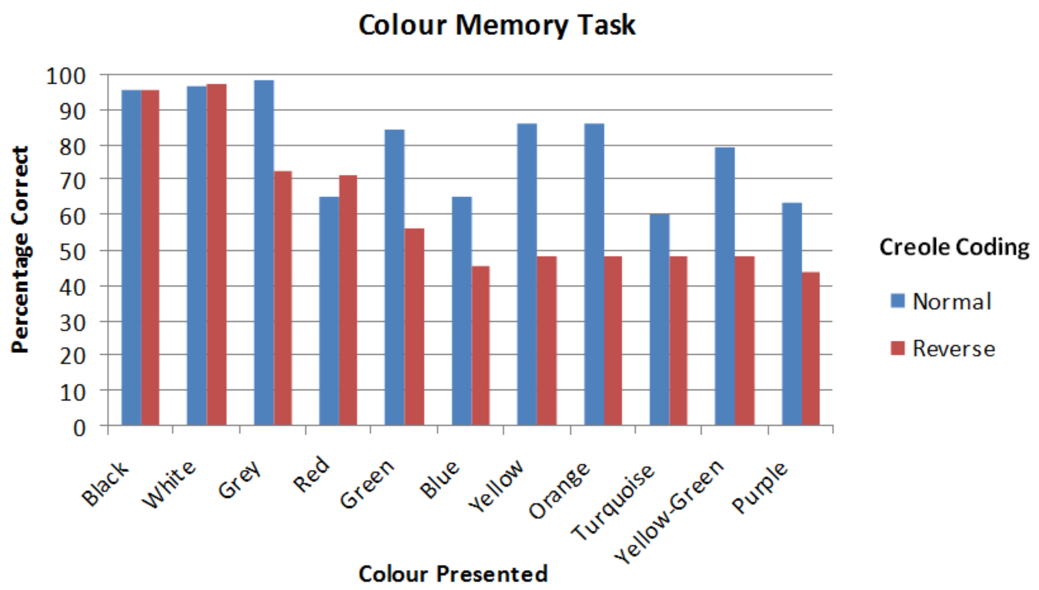


Figure 6.6. Cumulative scores for the 'normal' and 'reverse' colour mappings. ‘Normal’ groups showed an advantage in colour memory for most unique hues and all non-unique hues relative to the ‘reverse’ group.

From the cumulative colour scores, mappings that are the same for both Creole groups such as black and white have similar scores as would be expected. However when different auditory mappings are used for the remaining colours, a general trend for eight of the remaining nine colours shows higher scores for the 'normal' mapping over the 'reverse' mappings. The single exception to this is the 'red' coding where /i/ outperforms /u/ representations, despite /i/ also being an even better representation of green hues as seen in the 'normal' coding group. Despite 'red' being the only score that is lower for the 'normal' coding group, other more complex colours that utilise red (orange, purple) still show a substantial improvement using this red-/u/ mapping. It appears that colours that utilise the yellow coding of 3-high pitched tones have the highest improvement (yellow, orange, yellow-green), this is despite any increased luminance being coded separately as the highest pitched pure tone representing increased whiteness. The fact that 'normal' and 'reverse' groups' colour memory do not appear to differ on memorising the 'black' and 'white' colours (where they share the same colour-sound mappings), illustrates that the 'reverse' group was not globally worse, but that they are selectively worse when their auditory representations differ from the 'normal' group.

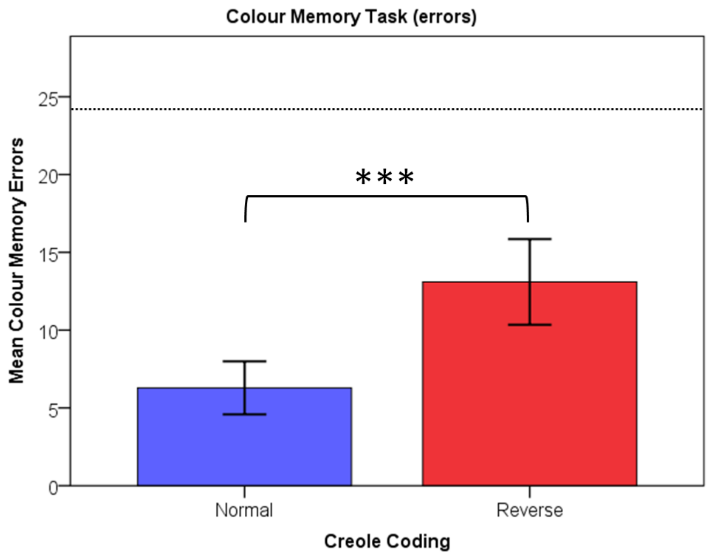


Figure 6.7. Mean colour memory task errors out of twenty-seven potential errors. Lower scores indicate better performance. The 'normal' Creole coding group made significantly fewer colour memory errors than the 'reverse' Creole coding group. Dotted line showed error rate expected by chance alone. Error bars indicate 95% confidence intervals. Key: *** = $p < .001$

For a comparison of overall memory scores, white and black were excluded as these were coded identically across Creole groups. A statistical analysis of colour memory errors revealed a significant main effect of group, $t\ (33.36) = -4.39$, $p < .001$, $d = -1.39$. The 'normal' Creole coding group made on average 6.29 errors (SD = 3.74) while the 'reverse' Creole coding group made significantly more errors on average at 13.10 (SD = 6.05) out of 27 potential errors (see fig. 6.7). This finding shows that participants who had the 'normal' Creole coding had an increased ability to effectively memorise and utilise their sound-colour mappings after brief training (<15 minutes with the device). In addition, all groups showed a greater than chance level (11.11% or 3 out of 27) of correctly identifying sounds as their correct colour representation.

6.4.2. Object identification task

In the object identification task, twenty trials were administered were participants were tasked with identifying a single fruit/vegetable presented on the Creole. Participants had twenty different stimuli choices to choose between making the chance level for this task is at 5% or 1 in 20.

Overall correct scores for the 'normal' and 'reverse' groups was not found to be statistically significant from one another, t (39) = 1.61, p = .115 with mean scores of 6.14 (SD = 2.63) and 4.65 (SD = 3.28) respectively (see fig 6.8). From this we can conclude that overall object identification abilities are not affected by the particular sound-colour mappings used by each group.

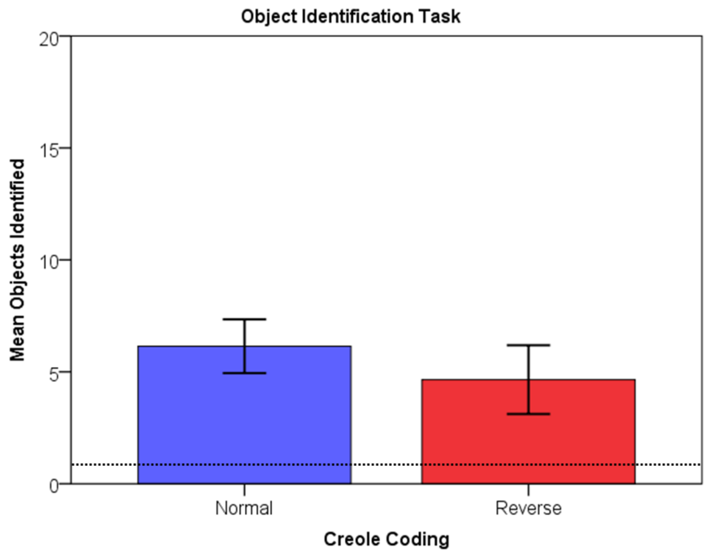


Figure 6.8. Graph showing the mean object identification task scores for both the 'normal' and 'reverse' Creole mapping groups. Higher scores indicate better performance. Dotted line shows chance level expected by random chance alone. Error bars indicate 95% confidence intervals.

While overall it appears that the auditory coding of the Creole has no effect on completing complex identification tasks, a further examination can breakdown the types of errors that participants make when making their selections. Two types of error can be examined from this, colour-errors and luminance errors. A colour-error is when a fruit/vegetable with a mismatched colour is selected, which can be assumed to be the result of hue misidentification. For example, if the stimuli 'bright *red* apple' is presented, the answer 'bright *green* lime' would have committed a colour-error, but the answer 'dark *red* pepper' would not commit a colour-error. A luminance error is when the lighting context is misidentified, if the target fruit is in the ‘bright’ lighting condition, any answer with ‘dark’ is considered a luminance-error (see fig 6.9).

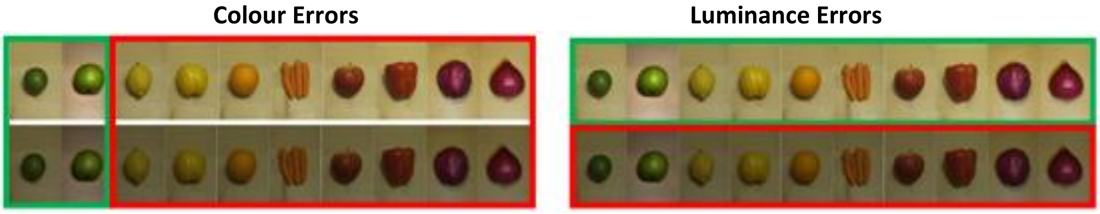


Figure 6.9. Example of types of error in the object identification task, if the image presented on the Creole was the 'bright green lime.' Left image shows the colour errors that could be made circled in red, any fruit/vegetables described as 'green' would not be a colour error, however any fruit/vegetable given a different descriptive colour term would be a 'colour error.' Similarly, luminance errors would only be when a 'dark' answer is given to a ‘bright’ stimulus or vice versa as seen circled in red.

A statistical analysis of the different types of error revealed a significant main effect of Creole group for the amount of colour-errors made during the object identification task, t (39) = -2.40, p = .021, d = -0.75. The 'normal' sound-colour coding had an average amount of 7.67 colour errors (SD = 2.92) made during the twenty trials, whereas the 'reverse' coding had a higher amount on average with 10.30 colour errors made (SD = 4.04). Both groups scored fewer errors than that expected by chance (16 errors), with the 'normal' group mistaking the colour of the target object less than those with the 'reverse' coding, showing that the sound-colour pairings do have a performance impact when applied to difficult naturalistic tasks involving colour. Luminance errors showed a similar trend however this was found to be marginally significant, t (39) = -2.01, p = .051, d = -0.63. The 'normal' coding group made fewer luminance-errors on average with 3.57 (SD = 2.14) in comparison to the 'reverse' coding group who had 5.00 errors on average (SD = 2.41). Both groups scored fewer errors than that expected by chance (10 errors), once again the 'normal' group made (marginally significant) fewer luminance errors than the 'reverse' group, which is surprising as the white and black codings for the device remained consistent across the coding types.

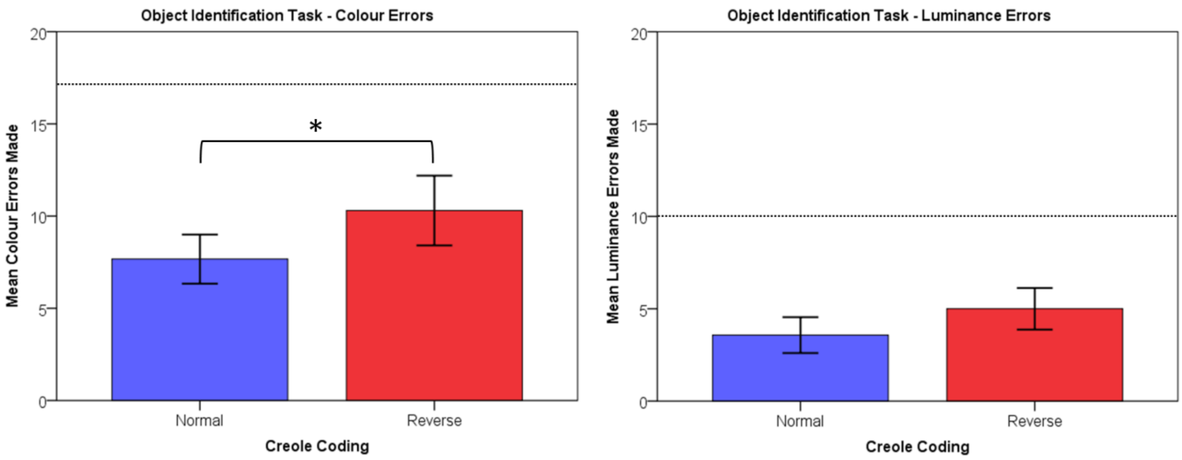


Figure 6.10. Left graph shows mean colour-errors made, right graph shows mean luminance-errors made during the object identification task. Lower scores indicate better performance. Dotted line shows chance level expected by random chance alone. Error bars indicate 95% confidence intervals. Key: * = p < .05.

It may be that the increased difficulty for identifying an object's colour for the reverse group may distract participants from accurately identifying the environmental luminance. Alternatively the background's yellowish hue may have had an impact on the identification of bright / dark conditions. Since yellow is easier to identify for the normal group as seen in the colour memory task (85.7% vs. 48.5%) it may be easier to understand variations of yellow. By accurately understanding the background, this could in turn allow an easier contrast with the colours of the target object. Alternatively the sounds themselves (C major or C minor chords) may differ in the user's ability to accurately contrast with the /i/ or /u/ vowel sounds represented through target objects featuring green or red hues. Supportive of this notion is that for the normal group the complex colours using C minor chords performed lower than those incorporating C major chords (60.3% and 63.5% vs. 85.7% and 79.4% respectively). During the colour memory task, the normal group mistook yellow (C major chord) for another hue only 1.6% of the time, however they mistook blue (C minor chord) for another hue 27% of the time. It is unclear whether this is an effect of C minor being harder to distinguish or not being the most intuitive representation for blue as compared to the C major chord which is a particularly effective representation of yellow.

Further exploration of the data associated with individual objects yielded some interesting results. By examining the percentage that each individual object was correctly identified (see figure 6.11), it was noted that the three easiest objects to discriminate were the same across groups, specifically the lime, orange and dark lemon (correct identification ranging between 60-40% of instances), but the most difficult varied according to group. This suggests that certain stimuli were easier to discriminate irrespective of their representation in sound, most likely through their shape and patterning information. The largest differences in performance between the normal and reverse groups primarily occurred for yellow and green objects (e.g. yellow pepper +26.2%, green apple +24.9%, dark green lime +16.4%, dark yellow lemon +10%), suggesting that these representations in sound had unique benefits for the normal group. The reverse group scored better than the normal group for certain objects containing red (e.g. dark red apple +15%, purple cabbage +15%) but the normal group also did better on different red objects (e.g. red apple +9.2%, red pepper +10.9%), so any benefit for red representations is far less stable. The results from the colour memory task (see figure 6.6) indicate that the normal group also originally had higher scores for yellow, green and blue colours but not red relative to the reverse group. Since there are no blue objects were presented any benefit cannot manifest here, so instead the normal groups' enhanced colour memory for yellow and green appears to enhance their discrimination for yellow and green coloured objects specifically.

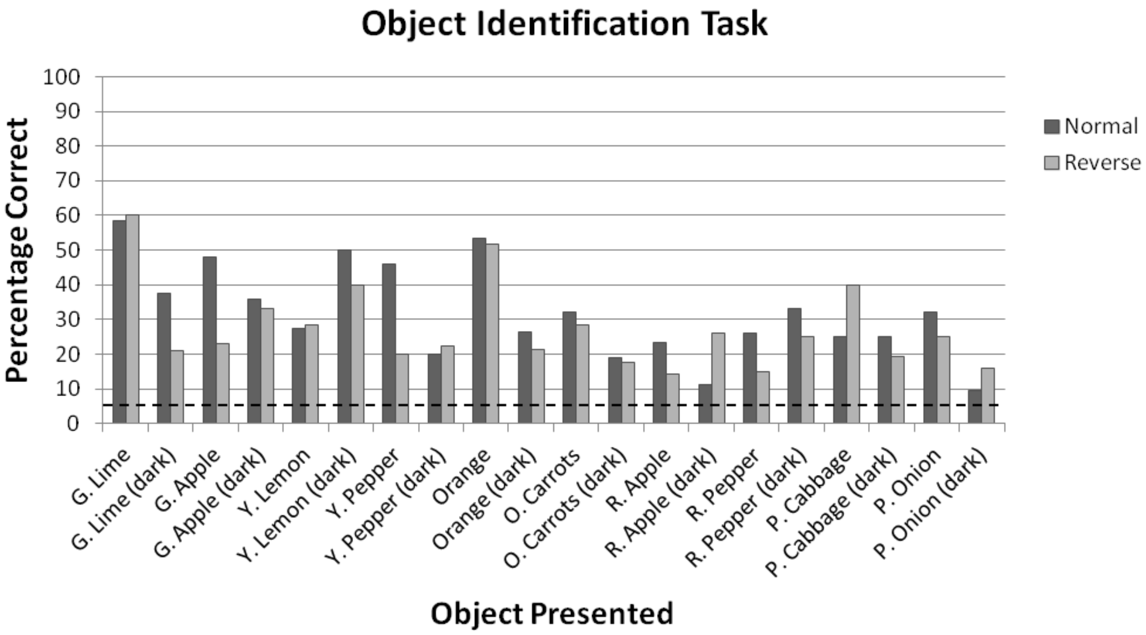


Figure 6.11. Percentage of individual objects identified by the 'normal' and 'reverse' colour mapping groups. Some objects show similar rates of identification across groups, whereas other objects demonstrate differences between what the 'normal' and 'reverse' colour coding groups can effectively discriminate. Dotted line shows score expected by random chance alone.

6.5. Discussion

The present experiment examines how the use of sound-colour correspondences affects the use of colour sensory substitution devices. We provide evidence that utilising sound-hue correspondences (Chapter five) improves memory for abstract computerised colours and that furthermore; this improved colour memory translates to naturalistic colour images to reduce colour-based errors. This finding underscores the idea that sounds are not arbitrary memory markers for visual dimensions such as hue or saturation. Instead, the evidence presented here suggests that specific combinations of sounds and colours can lead to optimal processing, aiding participants in bridging the modality-gap in sensory substitution. In addition, there was also a tendency to make fewer mistakes not directly related to sound-colour mappings (i.e. errors of luminance) and that while object identification tends to be improved this is not to a significant degree. The lack of a translation of fewer colour errors into more correct object answers may be due to other factors such as individual exploration styles. The findings have relevance for the use of sound-colour correspondences and future sensory substitution design, particularly for SSD users with previous visual experience.

6.5.1. Communicating colour through SSDs

Prior studies on colour SSDs have typically involved tasks for whom colour is the only route to a correct answer, such as with naming the colours of a flag (Ancuti et al., 2009), of matching otherwise identical coloured socks (Bologna et al., 2007) or of abstract computerised colour objects (Burch, 2012; Kahol, French, Bratton & Panchanathan, 2006). These tasks typically use the same colour categories as those coded into the device, for instance, a device that does not code 'orange' as a specific sound will not use orange in a formal behavioural task. This showcases that high colour resolutions allow for flexibility in understanding natural colours. Some more informal experiments have tried natural environments (Meers & Ward, 2004), or natural object identification (Maidenbaum et al., 2014). While these naturalistic approaches give merit to the utility of colour information, most devices substantially reduce the colour resolution given to users, making some objects under the same colour label indistinguishable. The ability to intuitively convey subtle variations in colour (e.g. red or orange) not only give additional cues to object identity, but how a stimulus might sound in darker and lighter environmental luminances as well. A high colour resolution can also convey the relationship between colours to those without veridical colour perception through the analogue between perceptual colour distance with that of perceptual auditory distance. This may act as a tool to teach colour concepts such as perceptual distance and mixing.

6.5.2. Correspondences and SSD design

The use of correspondences to convey the proprieties of visual information is not new, with initial audiovisual SSDs such as the vOICe (Meijer, 1992) utilising vertical, horizontal and luminance information in pitch, panning and loudness respectively (Collignon et al., 2009; Lewkowicz & Turkewitz, 1980; Melara & O'Brien, 1987; Roffler & Butler, 1967; Walker et al., 2010). However, the extent to which correspondences are the optimal choice for representing visual characteristics may need to consider the whole auditory soundscape. Brown, Macpherson and Ward (2011) report that the opposite mappings for luminance (i.e. black is mapped to loudness) may outperform correspondence-based mappings at target identification for dark objects on bright backgrounds. Part of the reason for this may be that the target's signal is salient relative to the background 'noise,' however of note is that colour judgements were not made by participants, so loud 'dark' objects may be intuitively thought of as 'bright' objects as per the loudness-luminance correspondence.

While the use of correspondences may not be the sole route to improving the understanding of SSD signal, the extent of the benefit of using correspondences over arbitrary mappings has not been the source of systematic investigation. As Chapter four's experiment shows, the addition of colour information is not necessarily of immediate benefit to users for object identification, especially when colour is only one of many cues that can be used for object identification. Instead, the auditory representation of the colour information has a significant impact on both colour replication, memory and identification (Chapter four, six). Even when a visual characteristic has multiple auditory correspondences (i.e. pitch-luminance and loudness-luminance) the use of specific correspondences can outperform others (Chapter four). Likewise the sound-hue correspondences investigated here show a largely consistent picture with one exception. Auditory representations of colour using correspondences (Chapter five) outperformed their opposites for colour memory scores, with the single exception of the /u/-red correspondence indicating that this correspondence may not be optimal. But even this reservation has a caveat since we also note /u/-red mappings work consistently well in complex colours that have a red component. Furthermore since /i/ appeared to be an effective communicator of greenness but the same could not be said for /u/ representing redness, one explanation is that participants could have used linguistic terms to help remember colour-sound associations, since /i/ is present in green, but /u/ is not in red. These points leave open future investigations into not only how multiple available correspondences provide optimal mappings for each visual dimension, both in isolation and together, but also how participants engage with these representations.

6.5.3. Why do correspondences improve performance?

The topic of how correspondences operate is currently the source of much debate and investigation, both with their neural underpinnings (Bien et al., 2012; Spence & Parise, 2012), innateness (Deroy & Spence, 2013b; Walker et al., 2010), relation to experience (Parise, Knorre & Ernst, 2014; Spence & Deroy, 2012) and level of automaticity (Spence & Deroy, 2013). Variations in a given correspondences' answer to these questions may be indicative of a particular 'type' of correspondence, from the lower-level structural and statistical correspondences to the higher level semantic variants (Spence, 2011). One potential interaction with SSDs is that certain correspondence types might be ideally suited to conveying visual information, or alternatively, that several of the same 'type' may influence each other when a single sound holds multiple pieces of visual information. Chapter four's finding that a loudness-luminance mapping outperforms a pitch-luminance mapping may be indicative of either optimal correspondences or an easier understanding of variations in loudness rather than pitch.

6.5.4. High colour resolutions in future SSDs

The use of high colour resolutions requires the use of multiple auditory dimensions, some of which are commonly used to denote spatial positions. As such the spatial distribution of colours may have to use alternative means to understand their location. One such possibility that remains unexplored in SSDs is that of head-related transfer function, whereby the modelling of how sounds change depending on where they are being emitted relative to the users' ear is replicated by use of filters (Blauert, 1997; Huopaniemi, Zacharov & Karjalainen, 1999). The intuitive understanding of these filtered sounds allows users to locate them in space, this in turn frees up additional auditory dimensions to allow for a higher auditory complexity as needed for higher colour resolutions.

6.5.5. Conclusion

This experiment demonstrates the first known work into investigating sound-hue correspondences for sensory substitution use. The show that specific sound-colour mappings do have an impact on performance through increasing colour memory, and reducing colour errors in new users of SSDs. Future studies should examine optimal representations of visual information in sound to further assist SSD users in overcoming the difficulty of crossing the 'modality-gap.'

7. Discussion

The aim of this thesis has been to explore how uniquely visual information such as colour can be represented through stimulation to alternative modalities, namely touch and hearing. This has been explored through the effect that variations of touch can have on visual photisms in synaesthesia, audio-visual matching preferences in the wider population and their influence on information processing through sensory substitution. Chapter two examines the shortcomings of greyscale SSDs, the utility of colour information and previous attempts to represent colour through touch and hearing in sensory substitution devices. The majority of previous psychological research into sensory substitution has principally been concerned with representing luminance and spatial information. The spatial resolution in particular for the end user is typically lower than that encoded by the device, limited by the end users' perceptual discrimination abilities and reaching functional resolutions up to but not exceeding the definitional criteria for blindness. Research into low-resolution vision has found colour information to help in the identification of objects and scenes, while the application of colour cues has been effectively utilised to segment objects from the environment in computer-vision studies (Crisman & Thorpe, 1993; Orwell, Remagnino & Jones, 2001; Torralba, 2009). It was argued that increases in colour information rather than spatial resolution may assist users more in object segmentation, discrimination and navigation. Recent evidence has suggested that novice greyscale SSD users reach a performance cap at 8 by 8 pixel resolutions for object discrimination using the vOICe (Brown, Simpson & Proulx, 2014). Whereas in scene identification, the addition of colour information nearly doubles the amount of correct identifications at 8 by 8 pixel resolutions, marking colour as an especially useful discriminating factor (Torralba, 2009). The utility of colour information in sensory substitution has been primarily explored in the engineering literature. These devices use a wide variety of approaches that differ in their spatial resolution, colour resolution and method of communicating colour to the SSD user. Despite these differences, colour SSDs have been used in a variety of studies examining colour knowledge, object identification and navigation in complex real-world environments. Comparisons with popular greyscale SSDs such as the vOICe have found that colour SSDs could be more resistant to changes in environmental illumination (Capalbo & Glenney, 2009), and that the added complexity of colour information does not necessarily entail a reduction in the functional spatial resolution (Ancuti, Ancuti & Bekaert, 2009). Instead, new evidence has emerged that colour (as represented by different timbres) may even help increase a users functional spatial resolution (Levy-Tzedek, Riemer & Amedi, 2014). These functional benefits of colour may be further improved using translations that are both perceptually distinctive dimensions and orientated to utilise intuitive cross-modal correspondences.

Expert late-blind users of SSDs have reported visual phenomenology from auditory SSD signals (Ward & Meijer, 2010), and while the same has not yet been reported for tactile-vision SSDs, tactile-orientation tasks have produced visual sensations in the blind (Ortiz et al., 2011). These reported experiences have been considered a form of acquired synaesthesia (Proulx, 2010; Ward & Wright, 2014). In order to explore how touch can be turned into visual phenomena, chapter three investigated developmental tactile-vision synaesthesia, exploring similarities in their reported phenomenology, psychophysics and tactile discrimination abilities. Due to the rarity of this variant of synaesthesia, there have been very few studies examining tactile-vision synaesthesia and so not a lot is known about common features present in this synaesthesia. To address this, a questionnaire was given to 21 respondents with self-reported tactile-vision synaesthesia. Several features reached high agreement between synaesthetes as inducing visual photisms, specifically emotional or sexual human touch and itchiness; notably all three of these have strong attentional and affective components. It was also reported that variations in texture were the strongest changers of visual content. All of these influences on the presence and quality of visual content have been previously associated with common underlying neural mechanisms such as the insula. Three tactile-vision synaesthetes that reported consistent touch-colour pairings were recruited for behavioural tests. Consistency tests were administered to establish genuineness which found that as a group they were more consistent than controls, however only one reached significance in isolation. It is suggested that currently employed consistency tests may not be optimal as touch-colour correspondences in controls may assisting controls too much (Ludwig & Simner, 2013). Based on the questionnaire data, alternatives are suggested to enhance sensitivity in confirming this synaesthesia by either looking for texture-colour consistency in objects that vary in multiple dimensions or consistency tests involving human touch for synaesthetes that do not experience photisms to inanimate objects. Similar to previously reported correspondences, a negative weight-luminance association was found for controls; however we newly report that this also influences synaesthetic photisms. This lends further weight to theories that posit specific mappings in synaesthesia are influenced by correspondences present in the wider population (Ludwig & Simner, 2013; Simner & Ludwig, 2012; Ward, Huckstep & Tsakanikos, 2006). When asked to draw their photisms in response to variations in touch stimulation, the synaesthetes showed a strong spatial mapping between the location and direction of stimulation on their hand with the location and direction of subsequent photisms. Spatial distances between points of stimulation, their orientation and relation to the synaesthetes' visual perspective all appear to be strongly represented in the distribution of colour in their photisms. This evidence suggests use of a shared spatial map between tactile and visual modalities. Tactile texture discrimination tests revealed that synaesthetes had a unique advantage in gauging distances between points of stimulation on the skin, but not for orientation or tactile-visual integration. These three tasks utilise different pathways, however the texture task has been reported to use the right angular gyrus, a location involved in tactile localisation on the body and crucial for transforming motor frames of reference into visual frames of reference. As such, these findings suggest regions likely recruited in tactile-vision synaesthesia and potential pathways that could explain the closely knit spatial mappings between touch and vision. These findings have implications for understanding how the tactile-visual system is altered when tactile stimulation provides visual phenomenology. This also provides parallels with how acquired tactile-visual synaesthesia may result following visual deprivation (Armel & Ramachandran, 1999), how tactile information may use visual areas to perceive (Ortiz et al., 2011) or process information from tactile-visual sensory substitution (Ptito, Moesgaard, Gjedde & Kupers, 2005).

Chapter four explores the representation of visual information in sound and touch during sensory substitution use. Unlike prior studies on colour SSDs, we compared the representation of colour (luminance, colour) and its transformation into sound (variations of pitch, loudness) using the same sensory substitution device, allowing the representations of spatial dimensions to remain the same. For object discrimination tasks, we found that greyscale codings that utilised variations of loudness outperformed those that used variations of pitch. Interestingly during colour replication tasks, both groups performed equivalently, suggesting an equal understanding of luminance information. Despite both of these codings using previously established correspondences, variations in loudness were easier for participants to use for naturalistic images. This indicates that the presentation of information to the user has a large impact on performance and that not all correspondences perform equivalently. For the colour codings we did not find that the mere presence of colour information assisted users over luminance for object discrimination, despite evidence suggesting this should be the case from low-resolution vision studies, computer-vision studies and previous colour SSD studies. Instead colour codings that also utilised luminance-loudness mappings performed equivalently with greyscale luminance-loudness mappings, suggesting that participants did not take advantage of this additional colour information. This was further borne out by the colour replication task, which suggested that the alternative hue-pitch colour coding had a better understanding of colour but a worse application to naturalistic images. The lack of engagement with available colour information suggests that the presentation of this information needs to be more intuitive to the user. It is suggested that more intuitive colour spaces such as those based on human perception and the use of correspondences for specific hues may allow users to engage with colour information more effectively.

Recently correspondences for specific hues and saturations have been reported for instrumental and vocal sounds. It is not known which auditory features drive these correlations and their influence on hue or saturation has not been convincingly disentangled from the influence of luminance. Chapter five explores a wide range of sounds and their influence on colour selections in equiluminant conditions. It was found that the formant structure of vowel sounds predicted saturation towards specific hues, but that when this structure was abstracted to complex sine wave sounds, different colour selection patterns emerged. From this it was determined that influences beyond the most dominant frequencies appear to have an effect. Furthermore, the context that these vowel sounds are in is important, with similar selection patterns being reapplied across several speakers in the same context. Investigations into more simple dimensions of sound revealed that saturation increases with pitch, loudness and harmonics. Pitch in particular is affected by context, so that relatively 'high' pitches are more saturated. Interestingly, saturation towards specific hues such as yellow are repeatedly observed when sounds feature frequencies over 800Hz, this was observed in simple and complex sine wave sounds as well as vocal sounds. Finally we show that many of these and previously observed luminance correspondences are maintained under the influence of LSD. This data provides a breakdown of why certain complex sounds are related to specific hues by looking at the influence of more fundamental auditory attributes and establishing the role of context in the application of these correspondences. Correspondences that are affected by context are likely to feature relative abstractions of 'high' and 'low' magnitudes, supporting theories such as AToM (Walsh, 2003). Whereas correspondences more resistant to changes in context such as the loudness of frequencies over 800Hz being correlated to yellow-saturation may use qualitatively different mechanisms (Palmer, Schloss, Xu & Prado-León, 2013; Spence, 2011). Future studies should look to examine the role of auditory context for correspondences and their effect on perceptual processing. This would further differentiate the underlying mechanisms of correspondences, in establishing whether relative magnitudes lead to slower incongruency effects or if this is the result of specific auditory features.

With the discovery of specific sounds being related to specific saturation and hues, Chapter six looks at the effect of applying this to information processing in sensory substitution. The Creole SSD was altered to reflect the findings of the previous chapter by utilising the sound-colour mappings found in CIE LUV colour space. These mappings were either congruent or incongruent with the correspondences previously observed for vowels and complex sine wave sounds. It was found that congruent mappings performed better on colour memory tasks, and for naturalistic object discrimination tasks, reduced not only colour-based errors but luminance-errors as well. This occurred despite the luminance mappings being identical across groups. It was concluded that the congruent use of correspondences aided colour memory and that this may result in less cognitive effort required to identify naturalistic colours, allowing users to allocate more cognitive effort for identifying other aspects of an image such as its lighting context. Improving the abilities of SSD users, particularly during early training phases may help engage with these devices and help adoption outside of the lab (Maidenbaum, Abboud & Amedi, 2014). In devices that encode high resolution colour information, blind individuals can not only have access to this visual information but can understand how colour operates, the perceptual distance between colours (Kahol, French, Bratton & Panchanathan, 2006) and share in their affective qualities (Palmer et al., 2013). SSDs also provide an increasingly cheap solution for visual rehabilitation. 80% of blindness occurs in developing countries where medical help may be inaccessible or too expensive to obtain (WHO, 2014), by contrast technologies capable of serving as SSDs are made continually cheaper making them a potential option for assisting rudimentary visual functioning and as 'visual' training prior to medical interventions (Maidenbaum et al., 2014). One potential use of SSDs is as neuro-rehabilitation by maintaining 'visual' areas through processing 'visual' information from SSDs prior to any potential medical solutions aimed at restoring vision itself (Reich, Maidenbaum & Amedi, 2012). Having visual areas continue to process visual information after acquired blindness may help preserve the function of these cortical regions to be better equipped to process veridical visual information should sight restoration be achievable. Whereas for the congenitally blind, having devices that can teach them the sensorimotor contingencies of vision may also better prepare them for any future sight recovery as well as provide practical benefits to navigation and object discrimination (Bach-y-Rita, 2002; Bologna, Deville & Pun, 2010; Maidenbaum, Arbel, Buchs, Shapira & Amedi, 2014; Ward & Wright, 2014). In order to increase the practicality and engagement of users with SSDs, continued research is required into establishing and understanding the optimal mappings for the efficient cross-modal transfer of visual information.

8. References

Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., & Amedi, A. (2014). EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. *Restorative Neurology and Neuroscience*, *32*(2), 247-257. doi: 10.3233/RNN-130338

Abraham, H. D. (1982). A chronic impairment of colour vision in users of LSD. *The British Journal of Psychiatry*, *140*(5), 518-520. doi: 10.1192/bjp.140.5.518

Abraham, H. D., & Wolf, E. (1988). Visual function in past users of LSD: Psychophysical findings. *Journal of Abnormal Psychology*, *97*(4), 443. doi: 10.1037/0021-843X.97.4.443

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, *7*(10), 1057-1058. doi: 10.1038/nn1312

Afra, P., Funke, M., & Matsuo, F. (2009). Acquired auditory-visual synesthesia: A window to early cross-modal sensory interactions. *Psychology Research and Behavior Management*, *2*, 31–37.

Aleman, A., Rutten, G. J. M., Sitskoorn, M. M., Dautzenberg, G., & Ramsey, N. F. (2001). Activation of striate cortex in the absence of visual stimulation: an fMRI study of synesthesia. *Neuroreport*, *12*(13), 2827-2830.

Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, *10*(6), 687-689. doi: 10.1038/nn1912

Ancuti, C., Ancuti, C., & Bekaert, P. (2009). ColEnViSon: Color Enhanced Visual Sonifier A Polyphonic Audio Texture and Salient Scene Analysis. *VISAPP 2009: Proceedings of the 4th International Conference on Computer Vision Theory and Applications*, 2, 566-572.

Anderson, H. P., Seth, A. K., Dienes, Z., & Ward, J. (2014). Can grapheme-color synesthesia be induced by hypnosis? *Frontiers in Human Neuroscience*, *8*, 220. doi: 10.3389/fnhum.2014.00220

Anema, H. A., Kessels, R. P. C., de Haan, E. H. F., Kappelle, L. J., Leijten, F. S., van Zandvoort, M. J. E., & Dijkerman, H. C. (2008). Differences in finger localisation performance of patients with finger agnosia. *Neuroreport*, *19*(14), 1429–1433. doi: 10.1097/WNR.0b013e32830e017b

Armel, K. C., & Ramachandran, V. S. (1999). Acquired synesthesia in retinitis pigmentosa. *Neurocase*, *5*(4), 293–296. doi: 10.1080/13554799908411982

Arno, P., Capelle, C., Wanet-Defalque, M. C., Catalan-Ahumada, M., & Veraart, C. (1999). Auditory coding of visual patterns for the blind. *Perception*, *28*, 1013-1030. doi: 10.1068/p2607

Arno, P., De Volder, A. G., Vanlierde, A., Wanet-Defalque, M. C., Streel, E., Robert, A., Sanabria-Bohórquez, S., & Veraart, C. (2001). Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage*, *13*(4), 632-645. doi: 10.1006/nimg.2000.0731

Asher, J. E., Aitken, M. R., Farooqi, N., Kurmani, S., & Baron-Cohen, S. (2006). Diagnosing and phenotyping visual synaesthesia: a preliminary evaluation of the revised test of genuineness (TOG-R). *Cortex*, *42*(2), 137-146. doi: 10.1016/S0010-9452(08)70337-X

Asher, J. E., Lamb, J. A., Brocklebank, D., Cazier, J. B., Maestrini, E., Addis, L., Sen, M., Baron-Cohen, S., & Monaco, A. P. (2009). A whole-genome scan and fine-mapping linkage study of auditory-visual synesthesia reveals evidence of linkage to chromosomes 2q24, 5q33, 6p12, and 12p12. *The American Journal of Human Genetics*, *84*(2), 279-285. doi: 10.1016/j.ajhg.2009.01.012

Auvray, M., Hanneton, S., Lenay, C., & O'Regan, K. (2005). There is something out there: distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, *4*(4), 505-521. doi: 10.1142/S0219635205001002

Auvray, M., Hanneton, S., & O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: Localisation and object recognition with 'The vOICe.' *Perception*, *36*(3), 416. doi: 10.1068/p5631

Ayama, M., & Ikeda, M. (1998). Brightness-to-luminance ratio of colored light in the entire chromaticity diagram. *Color Research & Application*, *23*(5), 274-287. doi: 10.1002/(SICI)1520-6378(199810)23:5<274::AID-COL4>3.0.CO;2-T

Bach-y-Rita, P. (1972). *Brain mechanisms in sensory substitution*. New York: Academic Press.

Bach-y-Rita, P. (1995). *Nonsynaptic diffusion neurotransmission and late brain reorganization*. New York: Demos-Vermande.

Bach-y-Rita, P. (2002). Sensory substitution and qualia. In A. Noë & E. Thompson (Eds.), *Vision and Mind* (pp. 497–514). Cambridge, MA: MIT Press.

Bach-y-Rita, P. (2004). Tactile sensory substitution studies. *Annals of the New York Academy of Sciences*, 1013: 83–91. doi: 10.1196/annals.1305.006

Bach-y-Rita, P., & W Kercel, S. (2003). Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*, *7*(12), 541-546. doi: 10.1016/j.tics.2003.10.013

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature*, *221*, 963-964. doi: 10.1038/221963a0

Bach-y-Rita, P., Kaczmarek, K., Tyler, M., & Garcia-Lara, J. (1998). Form perception with a 49-point electrotactile stimulus array on the tongue: A technical note. *Journal of Rehabilitation Research and Development*, 35, 427-430.

Banissy, M. J., Cohen Kadosh, R. C., Maus, G. W., Walsh, V., & Ward, J. (2009). Prevalence, characteristics and a neurocognitive model of mirror-touch synaesthesia. *Experimental Brain Research*, *198*(2-3), 261-272. doi: 10.1007/s00221-009-1810-9

Banissy, M. J., Stewart, L., Muggleton, N. G., Griffiths, T. D., Walsh, V. Y., Ward, J., & Kanai, R. (2012). Grapheme-color and tone-color synesthesia is associated with structural brain changes in visual regions implicated in color, form, and motion. *Cognitive Neuroscience*, *3*(1), 29–35. doi: 10.1080/17588928.2011.594499

Banissy, M. J., Tester, V., Muggleton, N. G., Janik, A. B., Davenport, A., Franklin, A., Walsh, V., & Ward, J. (2013). Synesthesia for color is linked to improved color perception but reduced motion perception. *Psychological Science*, *24*(12), 2390-2397. doi: 10.1177/0956797613492424

Banissy, M. J., Walsh, V., & Ward, J. (2009). Enhanced sensory perception in synaesthesia. *Experimental Brain Research*, *196*(4), 565–571. doi: 10.1007/s00221-009-1888-0

Bargary, G., & Mitchell, K. J. (2008). Synaesthesia and cortical connectivity.*Trends in Neurosciences*, *31*(7), 335-342. doi: 10.1016/j.tins.2008.03.007

Barnett, K. J., Finucane, C., Asher, J. E., Bargary, G., Corvin, A. P., Newell, F. N., & Mitchell, K. J. (2008a). Familial patterns and the origins of individual differences in synaesthesia. *Cognition*, *106*(2), 871-893. doi: 10.1016/j.cognition.2007.05.003

Barnett, K. J., Foxe, J. J., Molholm, S., Kelly, S. P., Shalgi, S., Mitchell, K. J., & Newell, F. N. (2008b). Differences in early sensory-perceptual processing in synesthesia: A visual evoked potential study. *Neuroimage*, *43*(3), 605–613. doi: 10.1016/j.neuroimage.2008.07.028

Baron-Cohen, S. (1996). Is there a normal phase of synaesthesia in development? *Psyche*, *2*(27), 223-228.

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colours: an experimental investigation of a case of synaesthesia. *Perception*, *16*(6), 761-767. doi: 10.1068/p160761

Baron-Cohen, S., Burtlf, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: prevalence and familiality. *Perception*, *25*, 1073-1079.

Barrass, S. (1997). *Auditory Information Design* (Doctoral dissertation, The Australian National University). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.6228&rep=rep1&type=pdf

Beeli, G., Esslen, M., & Jäncke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science*, *18*(9), 788-792. doi: 10.1111/j.1467-9280.2007.01980.x

Belkin, K., Martin, R., Kemp, S. E., & Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological Science*, *8*(4), 340-342. doi: 10.1111/j.1467-9280.1997.tb00450.x

Berk, T., Brownston, L., & Kaufman, A. (1982). A new color-naming system for graphics languages. *IEEE Computer Graphics and Applications*, *2*(3), 37-44. doi: 10.1109/MCG.1982.1674223

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study. *Neuroimage*, *59*(1), 663-672. doi: 10.1016/j.neuroimage.2011.06.095

Blake, R., Sobel, K. V., & James, T. W. (2004). Neural synergy between kinetic vision and touch. *Psychological Science*, *15*(6), 397–402. doi: 10.1111/j.0956-7976.2004.00691.x

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain*, *128*, 1571–1583. doi: 10.1093/brain/awh500

Blauert, J. (1997). *Spatial hearing. The psychophysics of human sound localization*. Cambridge, MA, USA: MIT Press.

Bleuler, E., & Lehmann, K. (1881). *Zwangsmässige Lichtempfindungen durch Schall und verwandte Erscheinungen auf dem Gebiete der anderen Sinnesempfindungen*. Leipzig: Fue's Verlag.

Bologna, G., Deville, B. & Pun, T. (2010). Blind navigation along a sinuous path by means of the See ColOr interface. In *Proceedings of the 3rd international Work-Conference on the interplay between Natural and Artifical Computation: Part II: Bioinspired Applications in Artifical and Natural Computation*, 235-243. doi: 10.1007/978-3-642-02267-8_26

Bologna, G., Deville, B., Pun, T., & Vinckenbosch, A. (2007). Transforming 3D coloured pixels into musical instrument notes for vision substitution applications. *Eurasip Journal on Image and Video Processing*. doi: 10.1155/2007/76204

Bologna, G., Deville, B., Vinckenbosch, M., & Pun, T. (2008). A perceptual interface for vision substitution in a color matching experiment. *2008 IEEE International Joint Conference on Neural Networks, Vols 1-8*, 1621–1628. doi: 10.1109/IJCNN.2008.4634014

Bond, B., & Stevens, S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception & Psychophysics*, *6*(6), 337–339. doi: 10.3758/BF03212787

Bornstein, M. (1987). Perceptual categories in vision audition, in S. Harnad (ed.), *Categorical Perception* (287-300), Cambridge: Cambridge University Press.

Boronkay, G. (2010). *Colour Conversion Centre*. Retrieved from http://ccc.orgfree.com/

Braaten, R. (1993). *Synesthetic correspondence between visual location and auditory pitch in infants*. Paper presented at the 34th Annual Meeting of the Psychonomic Society.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.

Brang, D., Hubbard, E. M., Coulson, S., Huang, M., & Ramachandran, V. S. (2010). Magnetoencephalography reveals early activation of V4 in grapheme-color synesthesia. *Neuroimage*, *53*(1), 268-274. doi: 10.1016/j.neuroimage.2010.06.008

Brang, D., & Ramachandran, V. S. (2008). Psychopharmacology of synesthesia; the role of serotonin S2a receptor activation. *Medical Hypotheses*, *70*(4), 903-904. doi: 10.1016/j.mehy.2007.09.007

Brang, D., Rouw, R., Ramachandran, V. S., & Coulson, S. (2011). Similarly shaped letters evoke similar colors in grapheme–color synesthesia. *Neuropsychologia*, *49*(5), 1355-1358. doi: 10.1016/j.neuropsychologia.2011.01.002

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, *48*(5), 630-637. doi: 10.1016/j.cortex.2011.06.008

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165-172. doi: 10.1016/j.cognition.2012.09.007

Brewer, C. A. (1999). Color use guidelines for data representation. *Proceedings of the Section on Statistical Graphics, American Statistical Association*, 55-60.

Brown, D., Macpherson, T., & Ward, J. (2011). Seeing with sound? Exploring different characteristics of a visual-to-auditory sensory substitution device. *Perception*, *40*, 1120-1135. doi: 10.1068/p6952

Brown, D. J., Simpson, A. J., & Proulx, M. J. (2014). Visual objects in the auditory system in sensory substitution: How much information do we need? *Multisensory Research*, *27*(5-6), 337-357. doi: 10.1163/22134808-00002462

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1831-1840. doi: 10.1098/rstb.2009.0028

Burch, D. (2012). *Development of a multiple contact haptic display with texture-enhanced graphics* (Doctoral dissertation, Virginia Commonwealth University). Retrieved from https://digarchive.library.vcu.edu/handle/10156/3836

Burch, D., & Pawluk, D. (2009). A cheap, portable haptic device for a method to relay 2-D texture-enriched graphical information to individuals who are visually impaired. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, 215-216. doi: 10.1145/1639642.1639682

Burton, H., MacLeod, A. M. K., Videen, T. O., & Raichle, M. E. (1997). Multiple foci in parietal and frontal cortex activated by rubbing embossed grating patterns across fingerpads: A positron emission tomography study in humans. *Cerebral Cortex*, *7*(1), 3–17. doi: 10.1093/cercor/7.1.3

Calvert, G., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. MIT press.

Capalbo, Z., & Glenney, B. (2009). Hearing color: Radical pluralistic realism and SSDs. *Proceedings of AP-CAP October 2009*, 135-140.

Capelle, C., Trullemans, C., Arno, P., & Veraart, C. (1998). A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *Biomedical Engineering, IEEE Transactions on*, *45*(10), 1279-1293. doi: 10.1109/10.720206

Cappelletti, L., Ferri, M., & Nicoletti, G. (1998). Vibrotactile colour rendering for the visually impaired within the VIDET project. In M. R. Stein (Ed.), *Telemanipulator and Telepresence Technologies V*, 3524, 92–96.

Chebat, D.-R., Rainville, C., Kupers, R., & Ptito, M. (2007). Tactile-'visual' acuity of the tongue in early blind individuals. *Neuroreport*, *18*(18), 1901–1904. doi: 10.1097/WNR.0b013e3282f2a63

Chebat, D.-R., Schneider, F., Kupers, R., & Ptito, M. (2011). Navigation with a sensory substitution device in congenitally blind individuals. *Neuroreport*, *22*(7), 342–347. doi: 10.1097/WNR.0b013e3283462def

Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, *41*(3), 339-353. doi: 10.1068/p7161

Cohen Kadosh, R., Henik, A., Catena, A., Walsh, V., & Fuentes, L. J. (2009). Induced cross-modal synaesthetic experience without abnormal neuronal connections. *Psychological Science*, *20*(2), 258-265. doi: 10.1111/j.1467-9280.2009.02286.x

Cohen Kadosh, R., Cohen Kadosh, K. C., & Henik, A. (2008). When brightness counts: The neuronal correlate of numerical–luminance interference. *Cerebral Cortex*, *18*(2), 337-343. doi: 10.1093/cercor/bhm058

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*(2), 132-147. doi: 10.1016/j.pneurobio.2007.11.001

Colizoli, O., Murre, J. M., & Rouw, R. (2012). Pseudo-synesthesia through reading books with colored letters. *PloS One*, *7*(6), e39799. doi: 10.1371/journal.pone.0039799

Collier, G. L. (1996). Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion*, *20*(1), 1-32. doi: 10.1007/BF02251005

Collignon, O., Lassonde, M., Lepore, F., Bastien, D., & Veraart, C. (2007). Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cerebral Cortex*, *17*(2), 457-465. doi: 10.1093/cercor/bhj162

Collins, M. (1929). A case of synaesthesia. *The Journal of General Psychology*, *2*(1), 12-27.

Coriat, I. H. (1913). A case of synesthesia. *The Journal of Abnormal Psychology*, *8*(1), 38.

Craig, J. C. (1999). Grating orientation as a measure of tactile spatial acuity. *Somatosensory and Motor Research*, *16*(3), 197–206. doi: 10.1080/08990229970456

Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*(8), 1196–1208. doi: 10.1016/S0028-3932(01)00224-X

Crisinel, A. S., & Spence, C. (2012). A fruity note: crossmodal associations between odors and musical notes. *Chemical Senses*, *37*(2), 151-158. doi: 10.1093/chemse/bjr085

Crisinel, A. S., Cosser, S., King, S., Jones, R., Petrie, J., & Spence, C. (2012). A bittersweet symphony: systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Quality and Preference*, *24*(1), 201-204.

Crisman, J. & Thorpe, C. (1993). SCARF: A Color Vision System that Tracks Roads and Intersections. *IEEE Transactions on Robotics and Automation*, 9(1), 49-58. doi: 10.1109/70.210794

Cronly–Dillon, J., Persaud, K. C., & Blore, R. (2000). Blind subjects construct conscious mental images of visual scenes encoded in musical form. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1458), 2231-2238. doi: 10.1098/rspb.2000.1273

Cronly-Dillon, J., Persaud, K., & Gregory, R. P. F. (1999). The perception of visual images encoded in musical form: a study in cross-modality information transfer. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1436), 2427-2433. doi: 10.1098/rspb.1999.0942

Cutsforth, T. D. (1925). The role of emotion in a synaesthetic subject. *American Journal of Psychology*, 36, 527–543.

Cytowic, R. E. (1989). *Synaesthesia: A union of the senses*. New York: Springer.

Cytowic, R. E., (1997). Synaesthesia: Phenomenology and neuropsychology—A review of current knowledge. In S. Baron-Cohen, & J. E. Harrison (Eds.), *Synaesthesia: Classic and Contemporary Readings* (1st ed., pp.17-39). Malden: Blackwell Publishing.

Cytowic, R. E., & Eagleman, D. (2009). *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*. MIT Press.

D'Andrade, R., & Egan, M. (1974). The colors of emotion. *American Ethnologist*, *1*(1), 49-63. doi: 10.1525/ae.1974.1.1.02a00030

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), 1049-1056. doi: 10.1038/79871

Dann, K. T. (1998). *Bright Colors Falsely Seen: Synaesthesia and the Search for Transcendental Knowledge*. Yale University Press.

Day, S. (2014). *Demographic Aspects of Synesthesia* [Fact sheet]. Retrieved from http://www.daysyn.com/Types-of-Syn.html

De Volder, A. G., Catalan-Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A., Michel, C., & Veraart, C. (1999). Changes in occipital cortex activity in early blind humans using a sensory substitution device. *Brain Research*, *826*(1), 128-134. doi: 10.1016/S0006-8993(99)01275-5

de Winter, J. C., & Dodou, D. (2010). Five-point Likert items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, *15*(11), 1-12.

Demattè, M. L., Sanabria, D., Sugarman, R., & Spence, C. (2006). Cross-modal interactions between olfaction and touch. *Chemical Senses*, *31*(4), 291-300.

Deroy, O., & Spence, C. (2013a). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, *20*(4), 643-664. doi: 10.3758/s13423-013-0387-2

Deroy, O., & Spence, C. (2013b). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1240-1253. doi: 10.1016/j.neubiorev.2013.04.001

Deroy, O., & Valentin, D. (2011). Tasting liquid shapes: investigating the sensory basis of cross-modal correspondences. *Chemosensory Perception*, *4*(3), 80-90. doi: 10.1007/s12078-011-9097-1

Deshpande, G., Hu, X., Stilla, R., & Sathian, K. (2008). Effective connectivity during haptic perception: A study using Granger causality analysis of functional magnetic resonance imaging data. *Neuroimage*, *40*(4), 1807–1814. doi: 10.1016/j.neuroimage.2008.01.044

Deville, B., Bologna, G., Vinckenbosch, M., & Pun, T. (2009). See color: Seeing colours with an orchestra. In D. Lalanne, & J. Kohlas (Eds.) *Human Machine Interaction* (pp. 251-279). Springer Berlin Heidelberg.

Dionne, J. K., Meehan, S. K., Legon, W., & Staines, W. R. (2010). crossmodal influences in somatosensory cortex: Interaction of vision and touch. *Human Brain Mapping*, *31*(1), 14–25. doi: 10.1002/hbm.20841

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector versus associator synaesthetes. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 335-343. doi: 10.3758/CABN.4.3.335

Dixon, M. J., Smilek, D., Cudahy, C., & Merikle, P. M. (2000). Five plus two equals yellow. *Nature*, *406*(6794), 365-365. doi: 10.1038/35019148

Doel, K. (2003). SoundView: Sensing color images by kinesthetic audio. *Proceedings of the 2003 International Conference on Auditory Display, Boston, MA*.

Doel, K., Smilek, D., Bodnar, A., Chita, C., Corbett, R., Nekrasovski, D., & McGrenere, J. (2004). Geometric shape detection with SoundView. *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia*.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch psychophysical evidence for linguistic relativity. *Psychological Science*, *24*(5), 613–621. doi: 10.1177/0956797612457374

Dotov, D. G., Nie, L., & Chemero, A. (2010). A demonstration of the transition from ready-to-hand to unready-to-hand. *Plos One*, *5*(3). doi: 10.1371/journal.pone.0009433

Durette, B., Louveton, N., Alleysson, D., & Hérault, J. (2008). Visuo-auditory sensory substitution for mobility assistance: testing TheVIBE. In *Workshop on Computer Vision Applications for the Visually Impaired (2008)*.

Eagleman, D. M., & Goodale, M. A. (2009). Why color synesthesia involves more than color. *Trends in Cognitive Sciences*, *13*(7), 288-292. doi: 10.1016/j.tics.2009.03.009

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods*, *159*(1), 139-145. doi: 10.1016/j.jneumeth.2006.07.012

Ebert, J. (2005). Scientists with disabilities: Access all areas. *Nature*, 435, 552-554. doi: 10.1038/435552a

Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology*, *37*(05), 697-705.

Elias, L. J., Saucier, D. M., Hardie, C., & Sarty, G. E. (2003). Dissociating semantic and perceptual components of synaesthesia: behavioural and functional neuroanatomical investigations. *Cognitive Brain Research*, *16*(2), 232-237. doi: 10.1016/S0926-6410(02)00278-1

Else, L. (2012). Cyborg makes art using seventh sense. *New Scientist*, 215, 50. doi: 10.1016/S0262-4079(12)62082-X

Engel, S. (1999). Using neuroimaging to measure mental representations: Finding color-opponent neurons in visual cortex. *Current Directions in Psychological Science*, 8, 23-27. doi: 10.1111/1467-8721.00006

Esterman, M., Verstynen, T., Ivry, R. B., & Robertson, L. C. (2006). Coming unbound: Disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of Cognitive Neuroscience*, *18*(9), 1570–1576. doi: 10.1162/jocn.2006.18.9.1570

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1). doi: 10.1167/10.1.6

Fairchild, M. D. (1998). *Color Appearance Models*. Addison Wesley Longman.

Fairchild, M. D. (2005). *Color Appearance Models* (2nd ed.). Chichester, UK: John Wiley & Sons.

Fechner, G. T. (1871). Zur experimentalen Aesthetik.

Fernay, L., Reby, D., & Ward, J. (2012). Visualized voices: a case study of audio-visual synesthesia. *Neurocase*, *18*(1), 50-56. doi: 10.1080/13554794.2010.547863

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, *102*(2), 1213–1222. doi: 10.1121/1.421048

Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic–auditory associations in the "Bouba–Kiki" effect. *Cognition*, *132*(2), 164-173. doi: 10.1016/j.cognition.2014.03.015

Gal, D., Wheeler, S. C., & Shiv, B. (2011). *Cross-modal influences on gustatory perception*. Manuscript submitted for publication. Available at http://ssrn.com/abstract=1030197

Gallace, A., Boschin, E., & Spence, C. (2011). On the taste of "Bouba" and "Kiki": An exploration of word–food associations in neurologically normal participants. *Cognitive Neuroscience*, *2*(1), 34-46. doi: 10.1080/17588928.2010.516820

Galletti, C., Battaglini, P., & Fattori, P. (1991). Functional-properties of neurons in the anterior bank of the parietooccipital sulcus of the macaque monkey. *European Journal of Neuroscience*, *3*(5), 452–461. doi: 10.1111/j.1460-9568.1991.tb00832.x

Gebuis, T., Nijboer, T. C., & Van der Smagt, M. J. (2009). Multiple dimensions in bi-directional synesthesia. *European Journal of Neuroscience*, *29*(8), 1703-1710. doi: 10.1111/j.1460-9568.2009.06699.x

Gelfand, S. (2009). *Essentials of Audiology* (3rd ed.). Stuttgart: Thieme Medical Publishers.

Giannakis, K. (2001). *Sound mosaics: a graphical user interface for sound synthesis based on audio-visual associations* (Doctoral dissertation, Middlesex University). Retrieved from http://eprints.mdx.ac.uk/6634/1/Konstantinos-sound_mosaics.phd.pdf

Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: the color of smells. *The American Journal of Psychology*, 335-351. doi: 10.2307/1423010

Ginsberg, L. (1923). A case of synaesthesia. *The American Journal of Psychology*, 582-589. doi: 10.2307/1414059

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition*, *12*(6), 878–892. doi: 10.1080/13506280444000562

Golledge, R., Marston, J., Loomis, J., & Klatzky, R. (2004). Stated preferences for components of a personal guidance system for nonvisual navigation. *Journal of Visual Impairment & Blindness*, 98(3), 135-147.

Goller, A. I., Otten, L. J., & Ward, J. (2009). Seeing sounds and hearing colors: an event-related potential study of auditory–visual synesthesia. *Journal of Cognitive Neuroscience*, *21*(10), 1869-1881. doi: 10.1162/jocn.2009.21134

Goodyear, B. G., & Menon, R. S. (1998). Effect of luminance contrast on BOLD fMRI response in human primary visual areas. *Journal of Neurophysiology*, *79*(4), 2204–2207.

Grantham, D., Hornsby, B., & Erpenbeck, E. (2003). Auditory spatial resolution in horizontal, vertical, and diagonal planes. The *Journal of the Acoustical Society of America*, 114, 1009-1022. doi: 10.1121/1.1590970

Grey, J. (1975). *Exploration of Musical Timbre*. Stanford University, Dept. of Music Report STAN-M-2.

Grey, J. (1977). Multidimensional percept scaling of musical timbres. *Journal of the Acoustical Society of America*, *61*(5), 1270–1277. doi: 10.1121/1.381428

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Sciences*, *5*(1), 36-41. doi: 10.1016/S1364-6613(00)01571-0

Guarniero, G. (1974). Experience of tactile vision. *Perception*, *3*(1), 101-104. doi: 10.1068/p030101

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, *146*(2), 161–171. doi: 10.1007/s00221-002-1164-z

Guest, S., & Spence, C. (2003). What role does multisensory integration play in the visuotactile perception of texture? *International Journal of Psychophysiology*, *50* (1), 63-80. doi: 10.1016/S0167-8760(03)00125-9

Guild, J. (1931). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London: Series A, Containing Papers of a Mathematical or Physical Character*, 230: 149-187. doi: 10.1098/rsta.1932.0005

Gur, M., & Akri, V. (1992). Isoluminant stimuli may not expose the full contribution of color to visual functioning: Spatial contrast sensitivity measurements indicate interaction between color and luminance processing. *Vision Research*, *32*(7), 1253-1262. doi: 10.1016/0042-6989(92)90220-D

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see auditory encoding of visual temporal sequences. *Psychological Science*, *16*(3), 228-235. doi: 10.1111/j.0956-7976.2005.00808.x

Hagen, M. C., Franzen, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience*, *16*(5), 957–964. doi: 10.1046/j.1460-9568.2002.02139.x

Haigh, A., Brown, D., Meijer, P. & Proulx, M. (2013). How well do you see what you hear? The acuity of visual-to-auditory sensory substitution. *Frontiers in Psychology*, 4, 330. doi: 10.3389/fpsyg.2013.00330

Hamilton-Fletcher, G., & Ward, J. (2013). Representing colour through hearing and touch in sensory substitution devices. *Multisensory Research*, *26*(6), 503. doi: 10.1163/22134808-00002434

Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: a versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, *7*(4), 269-276. doi: 10.1080/11762322.2010.512734

Hansen, T., Walter, S., & Gegenfurtner, K. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, 7, 1-15. doi: 10.1167/7.4.2

Hanson-Vaux, G., Crisinel, A. S., & Spence, C. (2013). Smelling shapes: Crossmodal correspondences between odors and shapes. *Chemical Senses*, *38*(2), 161-166. doi: 10.1093/chemse/bjs087

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: a critical overview. In S. Harnad (Ed.), *Categorical Perception: The Groundwork of Cognition* (pp. 535-565). Cambridge University Press.

Harris, J., Arabzadeh, E., Fairhall, A., Benito, C., & Diamond, M. (2006). Factors affecting frequency discrimination of vibrotactile stimuli: implications for cortical encoding. *PLoS One*, 1, 1. doi: 10.1371/journal.pone.0000100

Hartman, A. M., & Hollister, L. E. (1963). Effect of mescaline, lysergic acid diethylamide and psilocybin on color perception. *Psychopharmacologia*, *4*(6), 441-451. doi: 10.1007/BF00403349

Herde, L., Forster, C., Strupf, M., & Handwerker, H. O. (2007). Itch induced by a novel method leads to limbic deactivations—a functional MRI study. *Journal of Neurophysiology*, *98*(4), 2347-2356. doi: 10.1152/jn.00475.2007

Hidaka, S., Teramoto, W., Keetels, M., & Vroomen, J. (2013). Effect of pitch–space correspondence on sound-induced visual motion perception. *Experimental Brain Research*, *231*(1), 117-126. doi: 10.1007/s00221-013-3674-2

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*(1), 101-117. doi: 10.3758/BF03192848

Holle, H., Warne, K., Seth, A. K., Critchley, H. D., & Ward, J. (2012). Neural basis of contagious itch and why some people are more prone to it. *Proceedings of the National Academy of Sciences*, *109*(48), 19816-19821. doi: 10.1073/pnas.1216160109

Hollins, M., Goble, A., Whitsel, B., & Tommerdahl, M. (1990). Time course and action spectrum of vibrotactile adaptation. *Somatosensory and Motor Research*, 7, 205-221. doi: 10.3109/08990229009144707

Howells, T. H. (1944). The experimental development of color-tone synesthesia. *Journal of Experimental Psychology*, *34*(2), 87. doi: 10.1037/h0054424

Hribar, V., & Pawluk, D. (2011). A tactile-thermal display for haptic exploration of virtual paintings. *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, 221-222. doi: 10.1145/2049536.2049577

Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *American Journal of Psychology*, *109*(2), 219. doi: 10.2307/1423274

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme-color synesthetes: brain-behavior correlations. *Neuron*, *45*(6), 975-985. doi: 10.1016/j.neuron.2005.02.008

Hubbard, E. M., Brang, D., & Ramachandran, V. S. (2011). The cross-activation theory at 10. *Journal of Neuropsychology*, *5*(2), 152-177. doi: 10.1111/j.1748-6653.2011.02014.x

Hubbard, E. M., Manohar, S., & Ramachandran, V. S. (2006). Contrast affects the strength of synesthetic colors. *Cortex*, *42*(2), 184-194. doi: 10.1016/S0010-9452(08)70343-5

Hunt, R. W. G., & Pointer, M. R. (2011). *Measuring Colour* (4th ed.). Chichester, UK: John Wiley & Sons.

Huopaniemi, J., Zacharov, N., & Karjalainen, M. (1999). Objective and subjective evaluation of head-related transfer function filter design. *Journal of the Audio Engineering Society*, *47*(4), 218-239.

Hurvich, L., & Jameson, D. (1955). Some quantitative aspects of an opponent-colors theory: II. Brightness, saturation and hue in normal and dichromatic vision. *Journal of the Optical Society of America*, 45, 602-616. doi: 10.1364/JOSA.45.000602

Hurvich, L., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, 64, 384-404. doi: 10.1037/h0041403

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A Guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Isbell, H. (1959). Comparison of the reactions induced by psilocybin and LSD-25 in man. *Psychopharmacologia*, *1*(1), 29-38. doi: 10.1007/BF00408109

ISO (2003). *Acoustics – Normal Equal-Loudness Level Contours.* International Organisation for Standardisation; Geneva, Switzerland.

Jacobs, G. H. (2009). Evolution of colour vision in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1531), 2957-2967. doi: 10.1098/rstb.2009.0039

Jacobson, H. (1951a). The informational capacity of the human eye. *Science*, *113*(2933), 292-293. doi: 10.1126/science.113.2933.292

Jacobson, H. (1951b). Information and the human ear. *The Journal of the Acoustical Society of America*, *23*, 463. doi: 10.1121/1.1906788

Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryuken, M., & Müller-Gärtner, H. W. (1998). Intensity coding of auditory stimuli: an fMRI study. *Neuropsychologia*, *36*(9), 875-883. doi: 10.1016/S0028-3932(98)00019-0

Jansson, G. (1983). Tactile guidance of movement. *International Journal of Neuroscience*, *19*(1-4), 37-46. doi: 10.3109/00207458309148644

Jewanski, J. (2010). Color-tone analogies: A systematic presentation of the principles of correspondence. In D. Daniels, S. Naumann and J. Thoben, (Eds.) *Audiovisuology: Compendium. A Multidisciplinary Survey of Audiovisual Culture* (pp. 338-347). Köln: König.

Jewanski, J., Day, S. A., & Ward, J. (2009). A colorful albino: The first documented case of synaesthesia, by Georg Tobias Ludwig Sachs in 1812. *Journal of the History of the Neurosciences*, *18*(3), 293-303. doi: 10.1080/09647040802431946

Jewanski, J., Simner, J., Day, S. A., & Ward, J. (2011). The development of a scientific understanding of synesthesia from early case studies (1849–1873). *Journal of the History of the Neurosciences*, *20*(4), 284-305. doi: 10.1080/0964704X.2010.528240

Johnson, K., & Phillips, J. (1981). Tactile spatial-resolution. I. 2-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology*, *46*(6), 1177–1191.

Jordan, D. S. (1917). The colors of letters. *Science*, *46*(1187), 311-312.

Kaczmarek, K. A., & Haase, S. J. (2003). Pattern identification and perceived stimulus quality as a function of stimulation waveform on a fingertip-scanned electrotactile display. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, *11*(1), 9-16. doi: 10.1109/TNSRE.2003.810421

Kaczmarek, K. A., Bach-y-Rita, P., Tompkins, W. J., & Webster, J. G. (1985). A tactile vision-substitution system for the blind: computer-controlled partial image sequencing. *Biomedical Engineering, IEEE Transactions on*, (8), 602-608. doi: 10.1109/TBME.1985.325599

Kaczmarek, K. A., Tyler, M. E., & Bach-y-Rita, P. (1997). Pattern identification on a fingertip-scanned electrotactile display. *Proceedings of the 19th Annual International Conference of the IEEE / EMBS*, *4*, 1694-1696. doi: 10.1109/IEMBS.1997.757047

Kaczmarek, K. A., Webster, J., Bach-y-Rita, P., & Tompkins, W. (1991). Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering*, *38*, 1-16. doi: 10.1109/10.68204

Kahol, K., French, J., Bratton, L., & Panchanathan, S. (2006). Learning and perceiving colors haptically. *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, 173-180. doi: 10.1145/1168987.1169017

Kaiser, J. T. (2014). *Using the brainport for independent travel and obstacle avoidance* (Doctoral dissertation, University of Pittsburgh). Retrieved from http://d-scholarship.pitt.edu/21003/1/KaiserJT_ETD2014-2.pdf

Kajimoto, H., Kawakami, N., & Tachi, S. (2004). Electro-tactile display with tactile primary color approach. *International Conference on Intelligent Robots and System*. Retrieved from http://midas.inrf.uci.edu/groups/bats/uploads/Xsense/Kajimoto_IROS_2004.pdf

Keetels, M., & Vroomen, J. (2011). No effect of synesthetic congruency on temporal ventriloquism. *Attention, Perception, & Psychophysics*, *73*(1), 209-218. doi: 10.3758/s13414-010-0019-0

Kelly, K., & Judd, D. (1976). *Color; Universal Language and Dictionary of Names*. Milton Keynes: Lightning Source UK ltd.

Kemp, S. E., & Gilbert, A. N. (1997). Odor intensity and color lightness are correlated sensory dimensions. *The American journal of psychology*, *110*(1), 35. doi: 10.2307/1423699

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews Neuroscience*, *11*(6), 417–428. doi: 10.1038/nrn2833

Khan, A., Moideen, F., Lopez, J., Khoo, W., & Zhu, Z. (2012). KinDectect: kinect detecting objects. *Proceedings of the 13th international conference on Computer Helping People with Special Needs – Volume Part II*, 588-595. doi: 10.1007/978-3-642-31534-3_86

Kim, J. K., & Zatorre, R. J. (2008). Generalized learning of visual-to-auditory substitution in sighted individuals. *Brain research*, *1242*, 263-275. doi: 10.1016/j.brainres.2008.06.038

Klapetek, A., Ngo, M. K., & Spence, C. (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, *74*(6), 1154-1167. doi: 10.3758/s13414-012-0317-9

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Köhler, W. (1947). *Gestalt psychology: An Introduction to New Concepts in Modern Psychology*. New York: Liveright.

Kokjer, K. (1987). The information capacity of the human fingertip. *IEEE Transactions on Systems, Man and Cybernetics*, *17*, 100-102. doi: 10.1109/TSMC.1987.289337

Komisaruk, B. R., Whipple, B., Crawford, A., Grimes, S., Liu, W. C., Kalnin, A., & Mosier, K. (2004). Brain activation during vaginocervical self-stimulation and orgasm in women with complete spinal cord injury: fMRI evidence of mediation by the vagus nerves. *Brain Research*, *1024*(1), 77-88. doi: 10.1016/j.brainres.2004.07.029

Kosslyn, S. M., Thompson, W. L., Costantini-Ferrando, M. F., Alpert, N. M., & Spiegel, D. (2000). Hypnotic visual illusion alters color processing in the brain. *American Journal of Psychiatry*, *157*(8), 1279-1284. doi: 10.1176/appi.ajp.157.8.1279

Kostoulas, A. (2013). *On Likert Scales, Ordinal data and Mean Values*. Retrieved from http://achilleaskostoulas.com/2013/02/13/on-likert-scales-ordinal-data-and-mean-values/

Krill, A. E., Alpert, H. J., & Ostfeld, A. M. (1963). Effects of a hallucinogenic agent in totally blind subjects. *Archives of Ophthalmology*, *69*(2), 180-185. doi: 10.1001/archopht.1963.00960040186009

Krill, A. E., Wieland, A. M., & Ostfeld, A. M. (1960). The effect of two hallucinogenic agents on human retinal function. *Archives of Ophthalmology*, *64*(5), 724-733. doi: 10.1001/archopht.1960.01840010726015

Kupers, R., Chebat, D. R., Madsen, K. H., Paulson, O. B., & Ptito, M. (2010). Neural correlates of virtual route recognition in congenital blindness. *Proceedings of the National Academy of Sciences*, *107*(28), 12716-12721. doi: 10.1073/pnas.1006199107

Kupers, R., Fumal, A., de Noordhout, A. M., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Sciences*, *103*(35), 13256-13260. doi: 10.1073/pnas.0602925103

Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention Perception & Psychophysics*, *71*(7), 1439–1459. doi: 10.3758/APP.71.7.1439

Lederman, S. J., Gati, J., Servos, P., & Wilson, D. (2001). fMRI-derived cortical maps for haptic shape, texture, and hardness. *Cognitive Brain Research*, *12*(2), 307-313. doi: 10.1016/S0926-6410(01)00041-6

Lederman, S. J., Thorne, G., & Jones, B. (1986). Perception of texture by vision and touch – multidimensionality and intersensory integration. *Journal of Experimental Psychology-Human Perception and Performance*, *12*(2), 169–180. doi: 10.1037//0096-1523.12.2.169

Lee, V. K., Nau, A. C., Laymon, C., Chan, K. C., Rosario, B. L., & Fisher, C. (2014). Successful tactile based visual sensory substitution use functions independently of visual pathway integrity. *Frontiers in Human Neuroscience*, *8*, 291. doi: 10.3389/fnhum.2014.00291

Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., & Amedi, A. (2012a). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. *Restorative Neurology and Neuroscience*, 31, 1-11. doi: 10.3233/RNN-2012-110219

Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E., & Amedi, A. (2012b). Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific Reports*, 2, 949. doi: 10.1038/srep00949.

Levy-Tzedek, S., Riemer, D., & Amedi, A. (2014). Color improves "visual" acuity via sound. *Auditory Cognitive Neuroscience*, *8*, 358. doi: 10.3389/fnins.2014.00358

Lewkowicz, D., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology*, 6, 597-607. doi: 10.1037/0012-1649.16.6.597

Loomis, J. M., & Collins, C. C. (1978). Sensitivity to shifts of a point stimulus: An instance of tactile hyperacuity. *Perception & Psychophysics*, *24*(6), 487-492. doi: 10.3758/BF03198771

Lord, F. M. (1953). On the statistical treatment of football numbers. *The American Psychologist*, 8, 750-751. doi: 10.1037/h0063675

Lovero, K. L., Simmons, A. N., Aron, J. L., & Paulus, M. P. (2009). Anterior insular cortex anticipates impending stimulus significance. *Neuroimage*, *45*(3), 976-983. doi: 10.1016/j.neuroimage.2008.12.070

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences*, *108*(51), 20661–20665. doi: 10.1073/pnas.1112605108

Ludwig, V. U., & Simner, J. (2013). What colour does that *feel*? Tactile-visual mapping and the development of cross-modality. *Cortex,* 49, 1089-1099. doi: 10.1016/j.cortex.2012.04.004

Luke, D. P., & Terhune, D. B. (2013). The induction of synaesthesia with chemical agents: a systematic review. *Frontiers in Psychology*, *4*, 753. doi: 10.3389/fpsyg.2013.00753

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, *39*(12), 1304-1316. doi: 10.1016/S0028-3932(01)00119-1

MacDermott, M. M. (1940). *Vowel sounds in poetry: Their music and tone-colour*. London: Kegan Paul.

Maidenbaum, S., Abboud, S., & Amedi, A. (2014). Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews*, *41*, 3-15. doi: 10.1016/j.neubiorev.2013.11.007

Maidenbaum, S., Arbel, R., Buchs, G., Shapira, S., & Amedi, A. (2014). Vision through other senses: practical use of Sensory Substitution devices as assistive technology for visual rehabilitation. *Control and Automation (MED), 2014 22nd Mediterranean Conference of,* 182-187. doi: 10.1109/MED.2014.6961368

Mann, S., Huang, J., Janzen, R., Lo, R., Rampersad, V., Chen, A., & Doha, T. (2011). Blind navigation with a wearable range camera and vibrotactile helmet. *Proceedings of the 19th ACM international conference on Multimedia*, 1325-1328. doi: 10.1145/2072298.2072005

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, 87, 173-188.

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*, *82*(3), 303–331. doi: 10.1037/0033-2909.82.3.303

Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 384-394. doi: 10.1037/0096-1523.13.3.384

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, *28*(7), 903–923. doi: 10.1068/p2866

Martino, G., & Marks, L. E. (2000). Cross-modal interaction between vision and touch: the role of synesthetic correspondence. *Perception*, *29*(6), 745–754. doi: 10.1068/p2984

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, *10*(2), 61-65. doi: 10.1111/1467-8721.00116

Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., & Ptito, M. (2010). Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Research Bulletin*, 82, 264-270. doi: 10.1016/j.brainresbull.2010.05.001

Maurer, D., & Mondloch, C. J. (2005). Neonatal Synesthesia: A reevaluation. In L. C. Robertson & N. Sagiv (Eds.) *Synesthesia: Perspectives from Cognitive Neuroscience* (pp. 193-213). New York, NY, US: Oxford University Press.

McGlone, F., Vallbo, A. B., Olausson, H., Loken, L., & Wessberg, J. (2007). Discriminative touch and emotional touch. *Canadian Journal of Experimental Psychology*, *61*(3), 173–183. doi: 10.1037/cjep2007019

McGurk, H., & MacDonald. J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748. doi: 10.1038/264746a0

McMorrow, G., Wang, X., & Whelan, P. (1997). A color-to-speech sensory substitution device for the visually impaired. *Machine Vision Applications, Architectures, and Systems Integration VI 1997, Pittsburgh, USA*. doi: 10.1117/12.285572

Meers, S. & Ward, K. (2004). A vision system for providing 3D perception of the environment via transcutaneous electro-neural stimulation. *Proceedings of the Information Visualisation, 8th International Conference*, 546-552. doi: 10.1109/IV.2004.16

Meers, S. & Ward, K. (2005). A substitute vision system for providing 3d perception and gps navigation via electro-tactile stimulation. *Proceedings of the 1st International Conference on Sensing Technology, November 21-23, 2005, Palmerston North, New Zealand*.

Meers, S. & Ward, K. (2007). Substitute three-dimensional perception using depth and colour sensors. Retrieved from http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1585&context=infopapers

Meier, B., & Rothen, N. (2009). Training grapheme-colour associations produces a synaesthetic Stroop effect, but not a conditioned synaesthetic response. *Neuropsychologia*, *47*(4), 1208-1211. doi: 10.1016/j.neuropsychologia.2009.01.009

Meier, B., Rothen, N., & Walter, S. (2014). Developmental aspects of synaesthesia across the adult lifespan. *Frontiers in Human Neuroscience*, *8*, 129. doi: 10.3389/fnhum.2014.00129

Meijer, P. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39, 112-121. doi: 10.1109/10.121642

Meijer, P. (2014). *The vOICe Mini-Tutorial*. Retrieved from http://www.seeingwithsound.com/

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 69–79. doi: 10.1037/0096-1523.15.1.69

Melara, R. D., & O'Brien, T. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, 116, 323-336. doi: 10.1037/0096-3445.116.4.323

Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature*, *282*(5737), 403. doi: 10.1038/282403a0

Merabet, L. B., Battelli, L., Obretenova, S., Maguire, S., Meijer, P., & Pascual-Leone, A. (2009). Functional recruitment of visual cortex for sound encoded object identification in the blind. *Neuroreport*, *20*(2), 132-138. doi: 10.1097/WNR.0b013e32832104dc

Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., Kauffman, T., & Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. *Plos One*, *3*(8). doi: 10.1371/journal.pone.0003046

Merabet, L. B., Thut, G., Murray, B., Andrews, J., Hsiao, S., & Pascual-Leone, A. (2004). Feeling by sight or seeing by touch? *Neuron*, *42*(1), 173–179. doi: 10.1016/S0896-6273(04)00147-3

Merabet, L. B., Swisher, J. D., McMains, S. A., Halko, M. A., Amedi, A., Pascual-Leone, A., & Somers, D. C. (2007). Combined activation and deactivation of visual cortex during tactile sensory processing. *Journal of Neurophysiology*, *97*(2), 1633-1641. doi: 10.1152/jn.00806.2006

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1), 36-47. doi: 10.1016/j.heares.2006.05.004

Miletic, G., Hughes, B., & Bach-y-Rita, P. (1988). Vibrotactile stimulation: An educational program for spatial concept development. *Journal of Visual Impairment & Blindness*, *82*(9), 366-370.

Mills, C. B., Boteler, E. H., & Larcombe, G. K. (2003). "Seeing things in my head": A synesthete's images for music and notes. *Perception, 32*(11), 1359-1376. doi: 10.1068/p5100

Mills, C. B., Boteler, E. H., & Oliver, G. K. (1999). Digit synaesthesia: A case study using a Stroop-type test. *Cognitive Neuropsychology*, *16*(2), 181-191. doi: 10.1080/026432999380951

Miyahara, E. (2003). Focal colors and unique hues. *Perceptual and Motor Skills*, *97*(3), 1038-1042. doi: 10.2466/pms.2003.97.3f.1038

Miyahara, T., Koda, A., Sekiguchi, R. & Amemiya, T. (2012). A psychological experiment on the correspondence between colors and voiced vowels in non-synesthetes. *Kansei Engineering International Journal*, *11*(1), 27-34. doi: 10.5057/kei.11.27

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 133-136. doi: 10.3758/CABN.4.2.133

Moos, A., Simmons, D., Simner, J., & Smith, R. (2013). Color and texture associations in voice-induced synesthesia. *Frontiers in Psychology*, *4*, 568. doi: 10.3389/fpsyg.2013.00568

Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: Vowel colours in the ear of the beholder. *i-Perception*, *5*(2), 132-142. doi: 10.1068/i0626

Morgan, G., Goodson, F. & Jones, T. (1975). Age differences in the associations between felt temperatures and color choices. *The American Journal of Psychology*, 88(1), 125-130. doi: 10.2307/1421671

Morley, J., & Rowe, M. (1990). Perceived pitch of vibrotactile stimuli – effects of vibration amplitude, and implications for vibration frequency coding. *The Journal of Physiology*, *431*, 403–416. doi: 10.1113/jphysiol.1990.sp018336

Morrison, I., Björnsdotter, M., & Olausson, H. (2011). Vicarious responses to social touch in posterior insular cortex are tuned to pleasant caressing speeds. *The Journal of Neuroscience*, *31*(26), 9554-9562. doi: 10.1523/JNEUROSCI.0397-11.2011

Muggleton, N. G., Cowey, A., & Walsh, V. (2008). The role of the angular gyrus in visual conjunction search investigated using signal detection analysis and transcranial magnetic stimulation. *Neuropsychologia*, *46*(8), 2198-2202. doi: 10.1016/j.neuropsychologia.2008.02.016

Muggleton, N., Tsakanikos, E., Walsh, V., & Ward, J. (2007). Disruption of synaesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia*, *45*(7), 1582-1585. doi: 10.1016/j.neuropsychologia.2006.11.021

Mulert, C., Jäger, L., Propp, S., Karch, S., Störmann, S., Pogarell, O., Möller, H-J., Hegerl, U. (2005). Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI. *NeuroImage*, *28*(1), 49–58. doi: 10.1016/j.neuroimage.2005.05.041

Myers, C. S. (1911). A case of synaesthesia. *British Journal of Psychology, 1904-1920*, *4*(2), 228-238. doi: 10.1111/j.2044-8295.1911.tb00045.x

Myers, C. S. (1914). Two cases of synaesthesia. *British Journal of Psychology, 1904-1920*, *7*(1), 112-117. doi: 10.1111/j.2044-8295.1914.tb00246.x

Nayatani, Y. (1998). A colorimetric explanation of the Helmholtz–Kohlrausch effect. *Color Research & Application*, *23*(6), 374-378. doi: 10.1002/(SICI)1520-6378(199812)23:6<374::AID-COL5>3.0.CO;2-W

Neufeld, J., Sinke, C., Dillo, W., Emrich, H. M., Szycik, G. R., Dima, D., Bleich, S., & Zedler, M. (2012). The neural correlates of coloured music: A functional MRI investigation of auditory–visual synaesthesia. *Neuropsychologia*, *50*(1), 85-89. doi: 10.1016/j.neuropsychologia.2011.11.001

Neufeld, J., Sinke, C., Zedler, M., Emrich, H. M., & Szycik, G. R. (2012). Reduced audio–visual integration in synaesthetes indicated by the double-flash illusion. *Brain Research*, *1473*, 78-86. doi: 10.1016/j.brainres.2012.07.011

Newman, S. S. (1933). Further experiments in phonetic symbolism. *The American Journal of Psychology*, *45*, 53-75. doi: 10.2307/1414186

Newton, I. (1706). *Opticks: or a Treatise of the Reflections, Refractions, Inflections & Colours of Light*. New York: Courier Dover Publications.

Novich, S., Cheng, S., & Eagleman, D. M. (2011). Is synaesthesia one condition or many? A large-scale analysis reveals subgroups. *Journal of Neuropsychology*, *5*, 353–371. doi: 10.1111/j.1748-6653.2011.02015.x

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S., & Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience*, *5*(4), 371-375. doi: 10.1038/nn818

O'hare, J. J. (1956). Intersensory effects of visual stimuli on the minimum audible threshold. *The Journal of General Psychology*, *54*(2), 167-170. doi: 10.1080/00221309.1956.9920271

O'Mahony, M. (1983). Gustatory responses to nongustatory stimuli. Perception, 12, 627–633. doi: 10.1068/p120627

O'Regan, J. K. (2011). *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*. Oxford University Press.

O'Regan, K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioural and Brain Sciences*, 24, 939-1031.

Oberman, L. M., & Ramachandran, V. S. (2008). Preliminary evidence for deficits in multisensory integration in autism spectrum disorders: the mirror neuron hypothesis. *Social Neuroscience*, *3*(3-4), 348-355. doi: 10.1080/17470910701563681

Occelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013). The takete–maluma phenomenon in autism spectrum disorders. *Perception*, *42*(2), 233-241. doi: 10.1068/p7357

Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B. G., Starck, G., Ekholm, S., Strigo, I., Worslet, K., Vallbo, A. B., & Bushnell, M. C. (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience*, *5*(9), 900-904. doi: 10.1038/nn896

Orlandatou, K. (2012). The role of pitch and timbre in the synaesthetic experience. *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, Thessaloniki, Greece*.

Ortiz, T., Poch, J., Santos, J. M., Requena, C., Martínez, A. M., Ortiz-Terán, L., Turrero, J., Barcia, J., Nogales, R., Calvo, A., Martínez, J., Córdoba, J. & Pascual-Leone, A. (2011). Recruitment of occipital cortex during sensory substitution training linked to subjective experience of seeing in people with blindness. *PloS One*, *6*(8), e23264. doi: 10.1371/journal.pone.0023264

Orwell, J., Remagnino, P. M., & Jones, G. A. (2001). Optimal color quantization for real-time object recognition. *Real-Time Imaging*, *7*(5), 401–414. doi: 10.1006/rtim.2000.0209

Ou, L., Luo, M., Sun, P., Hu, N., Chen, H., Guan, S., Woodcock, A., Caivano, J., Huertas, R., Tremeau, A., Billger, M., Izaban, H. & Richter, K. (2010). A cross-cultural comparison of colour emotion fortwo-colour combinations. *Color Research and Application*, 37, 23-43. doi: 10.1002/col.20648

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, *114*(2), 173-186. doi: 10.1016/j.jecp.2012.05.004

Palmer, S. E., Schloss, K. B., Xu, Z. & Prado-León, L. R. (2013). Music-color associations are mediated by emotion. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1212562110

Parise, C. V., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, *442*(3), 257-261. doi: 10.1016/j.neulet.2008.07.010

Parise, C. V., & Spence, C. (2009). ‘When birds of a feather flock together’: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One*, *4*(5), e5664. doi: 10.1371/journal.pone.0005664

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, *111*(16), 6104-6108. doi: 10.1073/pnas.1322705111

Parke, F. (2003). *Color Spaces*. Retrieved from http://www.viz.tamu.edu/faculty/parke/ends375f03/notes/sec1_4.html

Pasqualotto, A., & Proulx, M. J. (2012). The role of visual experience for the neural basis of spatial cognition. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1179-1187. doi: 10.1016/j.neubiorev.2012.01.008

Pastore, R. E., Li, X. F., & Layer, J. K. (1990). Categorical perception of nonspeech chirps and bleats. *Perception & Psychophysics*, *48*(2), 151–156. doi: 10.3758/BF03207082

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437-442. doi: 10.1163/156856897X00366

Piesse, C. H. (1891). *Piesse's art of perfumery* (5th ed.). London: Piesse and Lubin.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983-993. doi: 10.1016/S0896-6273(04)00107-2

Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T., & Schnitzler, A. (2005). Image-to-sound conversion: experience-induced plasticity in auditory cortex of blindfolded adults. *Experimental Brain Research*, *167*(2), 287-291. doi: 10.1007/s00221-005-0060-8

Poynton, C. (1997). *Frequently Asked Questions about Color*. Retrieved from http://www.poynton.com/PDFs/ColorFAQ.pdf

Pridmore, R. W. (2007). Effects of luminance, wavelength and purity on the color attributes: Brief review with new data and perspectives. *Color Research & Application*, *32*(3), 208-222. doi: 10.1002/col.20312

Pritchard, J., Rothen, N., Coolbear, D., & Ward, J. (2013). Enhanced associative memory for colour (but not shape or location) in synaesthesia. *Cognition*, *127*(2), 230-234. doi: 10.1016/j.cognition.2012.12.012

Proulx, M. J. (2010). Synthetic synaesthesia and sensory substitution. *Consciousness and Cognition*, *19*(1), 501-503. doi: 10.1016/j.concog.2009.12.005

Proulx, M. J., Brown, D. J., Pasqualotto, A., & Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neuroscience & Biobehavioral Reviews*, *41*, 16-25.

Proulx, M. J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing 'where' through the ears: effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PLoS One*, *3* (3), e1840. doi: 10.1371/journal.pone.0001840

Ptito, M., Matteau, I., Gjedde, A., & Kupers, R. (2009). Recruitment of the middle temporal area by tactile motion in congenital blindness. *Neuroreport*, *20*(6), 543-547. doi: 10.1097/WNR.0b013e3283279909

Ptito, M., Moesgaard, S. M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain*, *128*(3), 606-614. doi: 10.1093/brain/awh380

Radvansky, G. A., Gibson, B. S., & McNerney, M. (2011). Synesthesia and memory: color congruency, von Restorff, and false memory effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 219. doi: 10.1037/a0021329

Ragert, P., Vandermeeren, Y., Camus, M., & Cohen, L. G. (2008). Improvement of spatial tactile acuity by transcranial direct current stimulation. *Clinical Neurophysiology*, *119*(4), 805–811. doi: 10.1016/j.clinph.2007.12.001

Raines, T. H. (1909). Report of a case of psychochromesthesia. *The Journal of Abnormal Psychology*, *4*(4), 249. doi: 10.1037/h0072856

Ramachandran, V. S., & Brang, D. (2008). Tactile-emotion synesthesia. *Neurocase*, *14*(5), 390-399. doi: 10.1080/13554790802363746Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia - a window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3-34.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, *288*(5), 52-59.

Ramachandran, V. S., Azoulai, S., Stone, L., Srinivasan, A. V., & Bijoy, N. (2005). Grasping metaphors and thinking with pictures: How brain damage might affect thought and language. Poster Presented at *the 12th Annual Meeting of the Cognitive Neuroscience Society, April 2005, New York, USA*.

Ramachandran, V. S., Miller, L., Livingstone, M. S., & Brang, D. (2012). Colored halos around faces and emotion-evoked colors: a new form of synesthesia. *Neurocase*, *18*(4), 352-358. doi: 10.1080/13554794.2011.608366

Regier, T., Kay, P., & Cook, R. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8386-8391. doi: 10.1073/pnas.0503281102

Reich, L., Maidenbaum, S., & Amedi, A. (2012). The brain as a flexible task machine: implications for visual rehabilitation using noninvasive vs. invasive approaches. *Current Opinion in Neurology*, *25*(1), 86-95. doi: 10.1097/WCO.0b013e32834ed723

Reichard, G. A., Jakobson, R., & Werth, E. (1949). Language and synesthesia. *Word*, *5*, 224-233.

Renier, L., Bruyer, R., & De Volder, A. G. (2006). Vertical-horizontal illusion present for sighted but not early blind humans using auditory substitution of vision. *Perception & psychophysics*, *68*(4), 535-542. doi: 10.3758/BF03208756

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C., & De Volder, A. G. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage*, *26*(2), 573-580. doi: 10.1016/j.neuroimage.2005.01.047

Ricciardi, E., Bonino, D., Pellegrini, S., & Pietrini, P. (2014). Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neuroscience & Biobehavioral Reviews*, *41*, 64-77. doi: 10.1016/j.neubiorev.2013.10.006

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition*, *98*(1), 53-84. doi: 10.1016/j.cognition.2004.11.003

Riggs, L. A., & Karwoski, T. (1934). Synaesthesia. *British Journal of Psychology, 25*(1), 29-41. doi: 10.1111/j.2044-8295.1934.tb00722.x

Rivest, J., & Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Research*, *36*(1), 53–66. doi: 10.1016/0042-6989(95)00056-6

Ro, T., Farnè, A., Johnson, R. M., Wedeen, V., Chu, Z., Wang, Z. J., Hunter, J. V., & Beauchamp, M. S. (2007). Feeling sounds after a thalamic lesion. *Annals of Neurology*, *62*(5), 433-441. doi: 10.1002/ana.21219

Robertson, A. R. (1977). The CIE 1976 color-difference formulae. *Color Research and Application*, 2, 7-11.

Roffler, S., & Butler, R. (1967). Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America*, 43, 1260-1266.

Rojczyk, A. (2011). Sound symbolism in vowels: Vowel quality, duration and pitch in sound-to-size correspondence. *Poznań Studies in Contemporary Linguistics PSiCL*, *47*, 602. doi: 10.2478/psicl-2011-0030

Root, R. T., & Ross, S. (1965). Further validation of subjective scales for loudness and brightness by means of cross-modality matching. *The American Journal of Psychology*, 285-289. doi: 10.2307/1420502

Rose, K. B. (1909). Some statistics on synaesthesia. *The American Journal of Psychology*, 447-447. doi: 10.2307/1413370

Rossi, J., Perales, F., Varona, J. & Roca, M. (2009). Col.diesis: transforming colour into melody and implementing the result in a colour sensor device. *Proceedings for the 2nd International Conference in Visualisation*. doi: 10.1109/VIZ.2009.49

Roth, P. Richoz, D., Petrucci, L., & Pun, T. (2001). An audio-haptic tool for non-visual image representation. *International Symposium on Signal Processing and its Applications (ISSPA)*, August 13th-16th, 2001. doi: 10.1109/ISSPA.2001.949776

Rothen, N., & Meier, B. (2010). Grapheme–colour synaesthesia yields an ordinary rather than extraordinary memory advantage: evidence from a group study. *Memory*, *18*(3), 258-264. doi: 10.1080/09658210903527308

Rothen, N., Nyffeler, T., von Wartburg, R., Müri, R., & Meier, B. (2010). Parieto-occipital suppression eliminates implicit bidirectionality in grapheme-colour synaesthesia. *Neuropsychologia*, *48*(12), 3482–3487. doi: 10.1016/j.neuropsychologia.2010.07.032

Rothen, N., Seth, A. K., Witzel, C., & Ward, J. (2013). Diagnosing synaesthesia with online colour pickers: maximising sensitivity and specificity. *Journal of Neuroscience Methods*, *215*(1), 156-160. doi: 10.1016/j.jneumeth.2013.02.009

Rothen, N., Tsakanikos, E., Meier, B., & Ward, J. (2013). Coloured Letters and Numbers (CLaN): a reliable factor-analysis based synaesthesia questionnaire.*Consciousness and cognition*, *22*(3), 1047-1060. doi: 10.1016/j.concog.2013.07.005

Rothen, N., Wantz, A. L., & Meier, B. (2011). Training synaesthesia. *Perception*, *40*, 1248-1250. doi: 10.1068/p6984

Rousselet, G., Joubert, O. & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition*, 12(6), 852-877. doi: 10.1080/13506280444000553

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*(6), 792-797. doi: 10.1038/nn1906

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *The Journal of Neuroscience*, *30*(18), 6205-6213. doi: 10.1523/JNEUROSCI.3444-09.2010

Rouw, R., Scholte, H. S., & Colizoli, O. (2011). Brain areas involved in synaesthesia: A review. *Journal of Neuropsychology*, *5*, 214–242. doi: 10.1111/j.1748-6653.2011.02006.x

Roy, E., & Hollins, M. (1998). A ratio code for vibrotactile pitch. *Somatosensory and Motor Research*, *15*(2), 134–145. doi: 10.1080/08990229870862

Rudmin, F., & Cappelli, M. (1983). Tone-taste synesthesia: A replication. *Perceptual and Motor Skills*, *56*(1), 118-118. doi: 10.2466/pms.1983.56.1.118

Saenz, M., & Koch, C. (2008). The sound of change: visually-induced auditory synesthesia. *Current Biology*, *18*(15), R650–R651. doi: 10.1016/j.cub.2008.06.014

Salvador, E., Cavallaro, A., & Ebrahimi, T. (2004). Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding*, *95*(2), 238-259. doi: 10.1016/j.cviu.2004.03.008

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225-239. doi: 10.1037/h0070931

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069-1078. doi: 10.1016/j.concog.2009.12.013

Sathian, K., & Stilla, R. (2010). Cross-modal plasticity of tactile perception in blindness. *Restorative Neurology and Neuroscience*, *28*(2), 271-281. doi: 10.3233/RNN-2010-0534

Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behavioural Brain Research*, *135*(1), 127-132. doi: 10.1016/S0166-4328(02)00141-9

Sathian, K., Zangaladze, A., Hoffman, J. M., & Grafton, S. T. (1997). Feeling with the mind's eye. *Neuroreport*, *8*(18), 3877–3881. doi: 10.1097/00001756-199712220-00008

Sampaio, E., Maris, S., & Bach-y-Rita, P. (2001). Brain plasticity: "visual" acuity of blind persons via the tongue. *Brain Research*, *908*(2), 204–207. doi: 10.1016/S0006-8993(01)02667-1

Schifferstein, H. N., & Tanudjaja, I. (2004). Visualising fragrances through colours: the mediating role of emotions. *Perception*, *33*(10), 1249-1266. doi: 10.1068/p5132

Schwarz, M. W., Cowan, W. B., & Beatty, J. C. (1987). An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Transactions on Graphics (TOG)*, *6*(2), 123-158. doi: 10.1145/31336.31338

Schwerdt, H. N., Tapson, J., & Etienne-Cummings, R. (2009). A color detection glove with haptic feedback for the visually disabled. *43rd Annual Conference on Information Science and Systems, 2009*, 681-686. doi: 10.1109/CISS.2009.5054804

Sebba, R. (1991). Structural correspondence between music and color. *Color Research & Application*, *16*(2), 81–88. doi: 10.1002/col.5080160206

Segond, H., Weiss, D., & Sampaio, E. (2005). Human spatial navigation via a visuo-tactile sensory substitution system. *Perception*, *34*(10), 1231–1249. doi: 10.1068/p3409

Seo, H. S., & Hummel, T. (2011). Auditory–olfactory integration: congruent or pleasant sounds amplify odor pleasantness. *Chemical Senses*, *36*(3), 301-309. doi: 10.1093/chemse/bjq129

Seo, H. S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., & Hummel, T. (2010). Cross-modal integration between odors and abstract symbols. *Neuroscience Letters*, *478*(3), 175-178. doi: 10.1016/j.neulet.2010.05.011

Shepard, R. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36, 2346-2353. doi: 10.1121/1.1919362

Shields, A. (2010). *The Development of Cross-modal Perception in Tactile and Visual Domains*. (Undergraduate Honours Dissertation, University of Edinbrugh) Retrieved from http://hdl.handle.net/1842/4950

Simner, J. (2007). Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends in cognitive sciences*, *11*(1), 23-29. doi: 10.1016/j.tics.2006.10.010

Simner, J. (2012). Defining synaesthesia. *British Journal of Psychology*, *103*, 1–15. doi: 10.1348/000712610X528305

Simner, J. (2013). The “rules” of synesthesia. In J. Simner and E. Hubbard (Eds.) *Oxford Handbook of Synesthesia*, (pp. 149-164). Oxford University Press.

Simner, J., & Bain, A. E. (2013). A longitudinal study of grapheme-color synesthesia in childhood: 6/7 years to 10/11 years. *Frontiers in Human Neuroscience*, *7*, 603. doi: 10.3389/fnhum.2013.00603

Simner, J., & Ludwig, V. U. (2012). The color of touch: A case of tactile-visual synaesthesia. *Neurocase*, *18*(2), 167–180. doi: 10.1080/13554794.2011.568503

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, *39*(4), 553. doi: 10.1068/p6591

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott, K., & Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033. doi: 10.1068/p5469

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. A. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology*, *22*(8), 1069–1085. doi: 10.1080/02643290500200122

Simpson, R. H., Quinn, M., & Ausubel, D. P. (1956). Synesthesia in children: Association of colors with pure tone frequencies. *The Journal of Genetic Psychology*, *89*(1), 95–103. doi: 10.1080/00221325.1956.10532990

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(8), 334–340. doi: 10.1016/j.tics.2009.05.001

Sinke, C., Halpern, J. H., Zedler, M., Neufeld, J., Emrich, H. M., & Passie, T. (2012). Genuine and drug-induced synesthesia: a comparison. *Consciousness and cognition*, *21*(3), 1419-1434. doi: 10.1016/j.concog.2012.03.009

Sjöström, C., & Rassmus-Gröhn, K. (1999). The sense of touch provides new computer interaction techniques for disabled people. *Technology and Disability*, 10, 45-52.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *Journal of Cognitive Neuroscience*, *13*(7), 930-936. doi: 10.1162/089892901753165845

Smilek, D., Dixon, M. J., & Merikle, P. M. (2005). Synaesthesia: Discordant male monozygotic twins. *Neurocase*, *11*(5), 363-370. doi: 10.1080/13554790500205413

Smilek, D., Moffatt, B. A., Pasternak, J., White, B. N., Dixon, M. D., & Merikle, P. M. (2002). Synaesthesia: A case study of discordant monozygotic twins. *Neurocase*, *8*(4), 338-342. doi: 10.1076/neur.8.3.338.16194

Smith, H. L. (1905). Synesthesia. *John Hopkins Hospital Bulletin*, *16*, 258-263.

Solomonow, M., Lyman, J., & Freedy, A. (1977). Electrotactile two-point discrimination as a function of frequency, body site, laterality, and stimulation codes. *Annals of Biomedical Engineering*, *5*(1), 47-60. doi: 10.1007/BF02409338

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics*, *73*(4), 971–995. doi: 10.3758/s13414-010-0073-7

Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned? *i-Perception*, *3*(5), 316. doi: 10.1068/i0526ic

Spence, C., & Deroy, O. (2013). How automatic are crossmodal correspondences? *Consciousness and Cognition*, *22*(1), 245-260. doi: 10.1016/j.concog.2012.12.006

Spence, C., & Driver, J. (2004). *Crossmodal Space and Crossmodal Attention.* Oxford University Press.

Spence, C., & Gallace, A. (2011). Tasting shapes and words. *Food Quality and Preference*, *22*(3), 290-295. doi: 10.1016/j.foodqual.2010.11.005

Spence, C., & Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i-Perception*, *3*(7), 410. doi: 10.1068/i0540ic

Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1298. doi: 10.1037/0096-1523.26.4.1298

Spence, C., Puccinelli, N. M., Grewal, D., & Roggeveen, A. L. (2014). Store atmospherics: A multisensory perspective. *Psychology & Marketing*, *31*(7), 472-488. doi: 10.1002/mar.20709

Spitoni, G. F., Pireddu, G., Cimmino, R. L., Galati, G., Priori, A., Lavidor, M., Jacobson, L., & Pizzamiglio, L. (2013). Right but not left angular gyrus modulates the metric component of the mental body representation: a tDCS study. *Experimental Brain Research*, *228*(1), 63-72. doi: 10.1007/s00221-013-3538-9

Steen, C. (2001). Visions shared: A firsthand look into synesthesia and art. *Leonardo*, *34*(3), 203-208. doi: 10.1162/002409401750286949

Steven, M. S., & Blakemore, C. (2004). Visual synaesthesia in the blind. *Perception*, *33*(7), 855-868. doi: 10.1068/p5160

Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences of the United States of America*, *54*(2), 407–411.

Stilla, R., & Sathian, K. (2008). Selective visuo-haptic processing of shape and texture. *Human Brain Mapping*, *29*(10), 1123–1138. doi: 10.1002/hbm.20456

Striem-Amit, E., & Amedi, A. (2014). Visual cortex extrastriate body-selective area activation in congenitally blind people “seeing” by using sounds. *Current Biology*, *24*(6), 687-692. doi: 10.1016/j.cub.2014.02.010

Striem-Amit, E., Cohen, L., Dehaene, S., & Amedi, A. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron*, *76*(3), 640-652. doi: 10.1016/j.neuron.2012.08.026

Striem-Amit, E., Guendelman, M., & Amedi, A. (2012). “Visual” acuity of the congenitally blind using visual-to-auditory sensory substitution. *Plos One*, *7*(3). doi: 10.1371/journal.pone.0033136

Syrkin, G., & Gur, M. (1997). Colour and luminance interact to improve pattern recognition. *Perception*, *26*, 127-140. doi: 10.1068/p260127

Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology*, *91*(3), 1282-1296. doi: 10.1152/jn.01125.2002

Tapson, J., Gurari, N., Diaz, J., Chicca, E., Sander, D., Pouliquen, P., & Etienne-Cummings, R. (2008). The feeling of color: A haptic feedback device for the visually disabled. *Proceedings for the IEEE Biomedical Circuits and Systems Conference, 2008. BioCAS*, 281-384. doi: BIOCAS.2008.4696954

Taylor, B. (1977). Dimensional interactions in vibrotactile information-processing. *Perception & Psychophysics*, *21*(5), 477–481. doi: 10.3758/BF03199505

Tegenthoff, M., Ragert, P., Pleger, B., Schwenkreis, P., Förster, A. F., Nicolas, V., & Dinse, H. R. (2005). Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 Hz rTMS. *PLoS Biology*, *3*(11), e362. doi: 10.1371/journal.pbio.0030362

Terhune, D. B., Tai, S., Cowey, A., Popescu, T., & Kadosh, R. C. (2011). Enhanced cortical excitability in grapheme-color synesthesia and its modulation. *Current Biology*, *21*(23), 2006–2009. doi: 10.1016/j.cub.2011.10.032

Terwogt, M. M., & Hoeksma, J. B. (1995). Colors and emotions: Preferences and combinations. *The Journal of General Psychology*, *122*(1), 5-17. doi: 10.1080/00221309.1995.9921217

Thornley Head, P. D. (2006). Synaesthesia: Pitch-colour isomorphism in RGB-space? *Cortex*, *42*(2), 164-174. doi: 10.1016/S0010-9452(08)70341-1

Titze, I. R. (1994). *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice-Hall.

Tkalcic, M., & Tasic, J. F. (2003). Colour spaces - perceptual, historical and applicational background. *EUROCON 2003*, *Computer as a Tool*, *8*(1), 304-308. doi: 10.1109/EURCON.2003.1248032

Tomson, S. N., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D. M., Bray, M., Leal, S., & Eagleman, D. M. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behavioural Brain Research*, *223*(1), 48-52. doi: 10.1016/j.bbr.2011.03.071

Torralba, A. (2009). How many pixels make an image? *Visual Neuroscience*, *26*(1), 123–131. doi: 10.1017/S0952523808080930

Valet, M., Pfab, F., Sprenger, T., Wöller, A., Zimmer, C., Behrendt, H., Ring, J., Darsow, U.,& Tölle, T. R. (2008). Cerebral processing of histamine-induced itch using short-term alternating temperature modulation–an FMRI study. *Journal of Investigative Dermatology*, *128*(2), 426-433. doi: 10.1038/sj.jid.5701002

Vallbo, A., & Johansson, R. (1984). Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Human Neurobiology*, *3*(1), 3–14.

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics*, *69*(5), 744-756. doi: 10.3758/BF03193776

Vatakis, A., Ghazanfar, A. A., & Spence, C. (2008). Facilitation of multisensory integration by the "unity effect" reveals that speech is special. *Journal of Vision*, *8*(9), 14. doi: 10.1167/8.9.14

Van Boven, R. W., & Johnson, K. O. (1994). The limit of tactile spatial resolution in humans grating orientation discrimination at the lip, tongue, and finger. *Neurology*, *44*(12), 2361-2361. doi: 10.1212/WNL.44.12.2361

Van Campen, C. (1999). Artistic and psychological experiments with synesthesia. *Leonardo*, *32*(1), 9-14. doi: 10.1162/002409499552948

Van Leeuwen, T. M., den Ouden, H. E., & Hagoort, P. (2011). Effective connectivity determines the nature of subjective experience in grapheme-color synesthesia. *The Journal of Neuroscience*, *31*(27), 9879-9884. doi: 10.1523/JNEUROSCI.0569-11.2011

Verrillo, R., Fraioli, A., & Smith, R. (1969). Sensation magnitude of vibrotactile stimuli. *Perception and Psychophysics*, 6, 366-372. doi: 10.3758/BF03212793

Visell, Y. (2009). Tactile sensory substitution: Models for enaction in HCI. *Interacting with Computers*, 21(1-2), 38-53. doi:10.1016/j.intcom.2008.08.004

Von Bismarck, G. (1974). Timbre of steady tones: A factorial investigation of its verbal attributes. *Acustica*, 30, 146-159.

Von Hornbostel, E. M. (1931). Über Geruchshelligkeit. *Pflügers Archive European Journal of Physiology*, *227*(1), 517–538. doi: 10.1007/BF01755351

Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, *42*(5), 491-502. doi: 10.3758/BF03209757

Walker, P. (2012). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception, & Psychophysics*, *74*(8), 1792–1809. doi: 10.3758/s13414-012-0341-9

Walker, P., & Smith, S. (1985). Stroop interference based on the multimodal correlates of haptic size and auditory pitch. *Perception*, *14*, 729–736. doi: 10.1068/p140729

Walker, P., Bremner, G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*, 21-25. Doi: 10.1177/0956797609354734

Walker, P., Francis, B. J., & Walker, L. (2010). The brightness-weight illusion darker objects look heavier but feel lighter. *Experimental Psychology*, *57*(6), 462–469. doi: 10.1027/1618-3169/a000057

Walker, L., Walker, P., & Francis, B. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception*, *41*(10), 1186. doi: 10.1068/p7149

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483-488. doi: 10.1016/j.tics.2003.09.002

Wan, C. Y., Wood, A. G., Reutens, D. C., & Wilson, S. J. (2010). Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344-348. doi: 10.1016/j.neuropsychologia.2009.08.016

Wang, D. & Brown, G. (2006). *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. Oxford: Wiley-Blackwell.

Ward, J. (2004). Emotionally mediated synaesthesia. *Cognitive Neuropsychology*, *21*(7), 761-772. doi: 10.1080/02643290342000393

Ward, J. (2013). Synesthesia. *Annual review of psychology*, *64*, 49-75. doi: 10.1146/annurev-psych-113011-143840

Ward, J., & Wright, T. (2014). Sensory substitution as an artificially acquired synaesthesia. *Neuroscience & Biobehavioural Reviews*, *41*, 26-35. doi: 10.1016/j.neubiorev.2012.07.007

Ward, J., & Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and cognition*, *19*(1), 492-500. doi: 10.1016/j.concog.2009.10.006

Ward, J., Banissy, M. J., & Jonas, C. N. (2008). Haptic perception and synaesthesia. In M Grünwald (Ed.), *Human Haptic Perception: Basics and Applications* (pp 259-265). Basel: Birkhäuser.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, *42*(2), 264–280. doi: 10.1016/S0010-9452(08)70352-6

Ward, J., Li, R., Salih, S., & Sagiv, N. (2007). Varieties of grapheme-colour synaesthesia: a new theory of phenomenological and behavioural differences. *Consciousness and Cognition*, *16*(4), 913-931. doi: 10.1016/j.concog.2006.09.012

Ward, J., Moore, S., Thompson-Lake, D., Salih, S., & Beck, B. (2008). The aesthetic appeal of auditory – visual synaesthetic perceptions in people without synaesthesia. *Perception*, *37*(8), 1285 – 1296. doi: 10.1068/p5815

Ward, J., & Sagiv, N. (2007). Synaesthesia for finger counting and dice patterns: A case of higher synaesthesia? *Neurocase*, *13*(2), 86-93. doi 10.1016/j.concog.2006.09.012

Way, T. P., & Barner, K. E. (1997a). Automatic visual to tactile translation--Part I: Human factors, access methods, and image manipulation. *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, *5*(1), 81–94. doi: 10.1109/86.559353

Way, T. P., & Barner, K. E. (1997b). Automatic visual to tactile translation--Part II: Evaluation of the TACTile Image Creation System. *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, *5*(1), 95–105. doi: 10.1109/86.559354

Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer Music Journal*, *3*(2), 45-52.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, *93*(1), 10-19. doi: 10.1016/j.bandl.2004.07.006

Whitaker, T. A., Simoes-Franklin, C., & Newell, F. N. (2008). Vision and touch: Independent or integrated systems for the perception of texture? *Brain Research*, *1242*, 59–72. doi: 10.1016/j.brainres.2008.05.037

WHO. (2014). *Visual Impairment and Blindness*. Retrieved from http://www.who.int/mediacentre/factsheets/fs282/en/index.html

Witthoft, N., & Winawer, J. (2013). Learning, memory, and synesthesia. *Psychological Science*, *24*(3), 258. doi: 10.1177/0956797612452573

Wrembel, M. (2009). On hearing colours - cross-modal associations in vowel perception in a non-synaesthetic population. *Poznań Studies in Contemporary Linguistics*, *45*(4), 595–612. doi: 10.2478/v10010-009-0028-0

Wright, W. D. (1928). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30, 4, 141-164. doi: 10.1088/1475-4878/33/3/301

Wright, W. D. (1941). The sensitivity of the eye to small colour differences. *Proceedings of the Physical Society*, 53, 93. doi: 10.1088/0959-5309/53/2/301

Yang, G.-H., Kyung, K.-U., Srinivasan, M. A., & Kwon, D.-S. (2006). Quantitative tactile display device with pin-array type tactile feedback and thermal feedback. *2006 IEEE International Conference on Robotics and Automation (icra), Vols 1-10,* 3917–3922. doi: 10.1109/ROBOT.2006.1642302

Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: What is superior about the memory of synaesthetes? *Quarterly Journal of Experimental Psychology*, *60*(5), 681–695. doi: 10.1080/17470210600785208

Yon, D., & Press, C. (2014). Back to the future: synaesthesia could be due to associative learning. *Frontiers in Psychology*, *5*, 702. doi: 10.3389/fpsyg.2014.00702

Zacharakis, A., Pastiadis, K., Papadelis, G. & Reiss, J. (2011). An investigation of musical timbre: uncovering salient semantic descriptors and perceptual dimensions. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 807-812.

Zamm, A., Schlaug, G., Eagleman, D. M., & Loui, P. (2013). Pathways to seeing music: Enhanced structural connectivity in colored-music synesthesia. *Neuroimage*, *74*, 359-366. doi: 10.1016/j.neuroimage.2013.02.024

Zangaladze, A., Epstein, C. M., Grafton, S. T., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, *401*(6753), 587–590. doi: 10.1038/44139

Zhang, M. M., Mariola, E., Stilla, R., Stoesz, M., Mao, H., Hu, X. P., & Sathian, K. (2005). Tactile discrimination of grating orientation: fMRI activation patterns. *Human Brain Mapping*, *25*(4), 370–377. doi: 10.1002/hbm.20107