Models of Colour Semiotics

Maryam Mohammadzadeh Darrodi

Submitted in accordance with the requirements for the degree of PhD

The University of Leeds

School of Design

September 2012

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgment

I wish to express my genuine appreciation to my first supervisor Professor Stephen Westland for his immense reserve of patience and knowledge. Also, I would like to thank my supervisor Doctor Kishore Budha for his devoted assistance and concern. This thesis would never have been completed without the inspiration and devotion of my Parents Professor Mohsen Mohammadzadeh and Behnaz Keybakhi. Moreover, I would like to thank my sister, the Ponsford family and my friends for being extremely supportive.

Additionally, please note that the graphical figures in this thesis have been designed and drawn by myself.

Abstract

This thesis is concerned with the relationship between colour and certain bi-polar characteristics. This topic has previously been studied through highly controlled laboratory-based experiments but this thesis starts from the question of whether experiments conducted in the laboratory (which are necessarily constrained to have low numbers of participants) can ever hope to capture the full complexity of the relationships being studied, since there are likely to be strong cultural and regional differences. The key advance in this work therefore was to explore the use of a web-based experiment for collecting data on a large scale and from all over the world. A laboratory-based experiment to explore colour semiotics was carried out and broadly supports the earlier work carried out by Ou *et al*. A novel paradigm for carrying out colour semiotic experiments based on a large-scale internet presentation and distributed over large numbers of participants (over 2000 from 58 countries) was then conceived and implemented. Comparison with the laboratory-based experiment broadly validated the use of this new paradigm. The large amount of data collected allowed an analysis of gender and cultural differences to be carried out and it was shown that cultural and age may be significant factors but that gender is probably not. The thesis has made a contribution in terms of collecting new data, generating new models, and testing a web-based paradigm for carrying out colour-based experiments. One application of the colour-semiotic models that has been developed at the end of this thesis is in the design process and a potential new software tool that could build a bridge between science and design has been considered.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

- ANOVA: Analysis of Variance
- AP: Active-passive
- C: Chroma
- CD: Clean-dirty
- CIE: Commission Internationale de l'Eclairage
- CMYK: Cyan-magenta-yellow-black
- Cons: Constant
- Cos: Cosine
- CRT: Cathode ray tube
- DF: Degree of freedom
- Email: Electronic mail
- Etc: And so on, Et cetera
- E.g.: For example
- FS: Fresh-stale
- GB: Great Britain
- GUI: Graphic user interface
- H: Hue
- HL: Heavy-light
- HP: Hewlett Packard Company
- HS: Hard-soft
- Max: Maximum
- MC: Modern-classical
- MF: Masculine-feminine
- Min: Minimum
- NCS: Natural colour system
- ICC: International Colour Consortium
- I.e.: That is
- Info: Information
- KMO: Kaiser Meyer Olkin
- L: Lightness
- LCD: Liquid crystal display
- LCH: Lightness Chroma Hue
- LD: Like-dislike
- RGB: Red-green-blue
- Sig: Significance
- Std: Standard deviation
- Tang: Tangent
- TR: Tense-relaxed
- UK: United Kingdom
- URL: Uniform resource locator
- US: United States of America
- Var: Variance
- WC: Warm-c

1 Introduction

Colours can evoke strong emotional responses in viewers and can also communicate meanings and or concepts through association. For example, in many western societies black is associated with death and the mourning process. Consequently, colour may play a role in imparting information, creating lasting identity and suggesting imagery and symbolic value. The literature reveals a confusing picture for a taxonomy of these effects and terms such as *colour semiotics*, *colour semantics* and *colour emotion* are variously, and sometimes inter-changeably, used. Semantics concerns the meanings of words whereas semiotics concerns the meanings of signs and symbols. In this thesis the term colour semiotics is used because it is the meaning imparted by the sensation that an observer experiences when they look at a colour with which this thesis is primarily concerned. For example, the *colour red* rather than any meaning that is imparted by the word *red* itself. The term colour emotion has emerged mainly from one or two research groups in the UK and Japan and would seem to be a somewhat restricted term. For example, we consider the effect of a reddish colour which is normally considered to be warm; does it make the observer feel warm or does it communicate warmth? Perhaps, in this case, both an emotional effect and a symbolic meaning occur. But if we consider colours that may be thought of as masculine; such a colour is unlikely to make the observer feel more male but rather may communicate or indicate maleness. It would seem therefore that the term colour semiotics is most appropriate for the work described in this thesis.

There seem to be at least three different origins for colour semiotics. Firstly there is the emotional or visceral impact of colours. Colours can have a strong emotional impact and can even affect our physiological state. For example, red colours have been cited to raise the blood pressure and colours have been reported to affect muscular strength. We fear the dark. Perhaps these effects are *innate* and have been present since the earliest days (the effect of red has sometimes been attributed to the colour of blood and our fear of black may relate to a primitive fear of the dark and unknown.) Secondly there are *socio-economic* origins. In western society purple became associated with wealth and royalty because purple dyestuff was more expensive than silver. Only extremely rich people could afford to wear purple and some organizations chose to use purple to make a statement about their wealth and power. Thirdly, some colours meanings are *cultural* in origin. The association of red with luck in China and the link between pink for girls and blue for boys in western society may originate in and be reinforced by cultural behaviour and shared understanding. For example, in the United Kingdom pink was associated with young boys until about 1920 after which blue came to signify the male professions, most notably the navy.

The importance of colour semiotics has been noted – most frequently in the cultural studies literature - in corporate visual identities, human computer interaction, political communication, and as a marker for gender and sexuality. However, although black, for example, is often associated with death it can have other meanings; it can be associated with power or evil, and the actual meaning in any particular situation depends upon the context in which the colour is used; it can also depend upon other aspects of visual appearance such as gloss and texture. Furthermore, the meanings for a colour can also depend upon culture and can vary over time. For example, in some countries black is not the colour that is most associated with death (white is used instead). The appropriate use of colour semiotics can impact greatly on the success of a design (particularly one that has a branding or marketing dimension). However, it is clear that colour meanings and associations can vary with a great many factors. On one hand the connection of meaning and colour seems obvious; on the other hand it seems idiosyncratic, unpredictable and anarchic. According to the cultural studies literature, social groups that share common purposes around colour are often relatively small and specialised compared to groups who share speech or visual communication. Some researchers go so far as to suggest that colour *per se* does not elicit response, but the particular meaning or significance of the colour is context-bound and varies from one person or situation to another. Despite the prevalence of this position in some fields, the colour science community has generally adopted a quite different position. Most robust scientific studies that have explored colour semiotics have done so for colour patches viewed in an abstract sense, devoid of context. It has been found that there is an effect of culture on the association of colours with certain words. This would seem to contradict greatly with the view that colour *per se* does not elicit response.

Nevertheless, most formal studies in the last decade have explored whether there are cultural, gender or age effects in terms of the meanings associated with colours by observers when viewing colours without context (typically square patches of colour viewed on a computer screen). For example, in one study observers from seven countries (Hong Kong, Japan, Thailand, Taiwan, Italy, Spain and Sweden) were asked to rate 214 colour samples each in terms of 12 bi-polar word pairs (e.g. soft-hard). The differences between the nationality groups were small despite the different cultural backgrounds. In another study 14 British and 17 Chinese observers assessed 20 colours in terms of 10 bi-polar word pairs [1]. The differences between the responses from the two groups were small with the exception of like-dislike and tense-relaxed.

It seems clear that colour *per se* does have meaning but the question of whether these meanings are consistent across culture, age and gender is not entirely clear. As Gage wrote, "To what extent different colours, such as red or black, have cross-cultural significance, is an altogether more difficult question." This is still an unanswered question and is the starting point for the work in this thesis.

Perhaps one reason why these formal studies have not been able to provide definitive answers to the question of whether colour meaning and emotion depends upon culture (and even gender) is because they have traditionally been carried out with quite small numbers of participants (typically 30 or less) because the experiments are carried out in laboratories using carefully controlled and calibrated equipment. One way to involve much greater numbers of participants is to use a web-based experiment and this is the approach that has been adopted in this thesis. Of course, web-based experiments have numerous disadvantages including potential sources of colour variation including, display technology, ambient illumination level, observer bias an, deficiencies and anomalies and operating software. However, the use of large-scale web-based experiments, although challenging, could potentially address questions that small-group studies could never robustly address and may provide some link between the seemingly opposing views of academics in colour science and cultural studies.

The research in this thesis starts from the question of whether experiments conducted in the laboratory (which are necessarily constrained to have low numbers of participants) can ever hope to capture the full complexity of the relationships being studied since there are likely to be strong cultural and regional differences. The key advance in this work therefore is to explore the use of a web-based experiment for collecting data on a large scale and from all over the world. The implementation of a web-based experiment in colour is subject to possible criticism. However, the colours displayed (and the viewing environment) will not be as well controlled as in a laboratory experiment and it is not easy to screen-out colour-defective observers. Nevertheless, there are examples of successful web-based colour experiments, despite the problems and the work in this thesis was in part inspired by the work by Moroney who has been carrying out a web-based colour-naming experiment for at least 6 years [2]. Therefore one of the main motivations of this research is to bridge the gap between design and colour science.

The main research questions of the thesis are as follows:

- Would a new laboratory-based experiment agree with the previous work carried out by other researchers? Are the derived models robust?
- What is the novel paradigm for carrying out colour semiotic experiments? How it can be implemented? How it can be validated?
- How can the effect of age, gender and cultural differences be investigated? And are they significant? How can they be modelled?
- How the data and models derived from colour semiotic experiments can be effectively presented to the community? How can this presentation effectively motivate the application of these models?

In summary, Chapters 2 and 3 form a body of literature review for the thesis followed by a thorough review of colour semiotics in Chapter 4. In Chapter 5 a laboratory-version of the web-based experiment is carried out. This is to enable a comparison between laboratory- and web-based paradigms and also to be able to relate the work in this thesis with the seminal work of Ou and colleagues [1]. Chapter 6 addresses the inverse problem for models that predict bi-polar characteristics from colour attributes. Alternative numerical method (adaptive gradient descent optimisation) is discussed and implemented to solve the inverse problem. The global online colour survey is described in Chapter 7 followed by the analysis of the data collected from the Global Online Survey. In this chapter, new models are developed and compared with previous research and also the laboratory based experiment. In Chapter 8 a summary of the work is addressed along with the contribution made in this thesis. Finally, implementation of one potential application of models that is the development of a colour selector tool for the design community is discussed which leads to an insight of possible future work.

2 Fundamentals of Colour

2.1 Introduction

The colour of an object makes it to be recognised and sensed easily. However, despite the fact that colour is perceived as a simple everyday phenomenon; there are deeper aspects to it. As an extreme example, colour-blind people do not see colours in the full spectrum and blind people have no visionary sense at all; but some of them feel colour and have emotions as they hear about it[1], so it can be said that colour is far more than what we see.

But talking about the visual aspect of colour, it is interesting how it is viewed by the eyes and recognised by the human brain. Substantial research has taken place on colour properties by colour scientists; in physics, colour is defined as "waves of light that differ in frequency and energy through the spectrum from red (low frequency and energy) to violet (high frequency and energy)" [3]. In section 2.3 of this chapter a full description of the human vision process and colour physics is explained. We also need to know how to communicate colour. The huge range of colours across the world is far too many to be identified by human linguistics alone. So there is a necessity for standard colour systems which take the colours into account. Of particular importance is a numerical system introduced by the *Commission Internationale de l'Eclairage* (CIE) [4, 5].

Initially when we have become familiar with colour properties, we can take the next step towards colour application where we can apply colour to objects of interest depending on their characteristics.

2.2 Colour terminology and colour order systems

Almost everyone knows what colour is, since it is the one of the very first concepts learnt in childhood. Colours are described in different ways around the world which can be caused by diversity in language, culture and even geographical locations. However, it is not easy to define colour. The reason could be that the huge number of colours which exist in the visual world, cannot be simply categorised in to a few representative groups for example red, green, yellow and blue as lots of information about that specific colour is lost; normal people often remark it as different shades of colour by using words such as warmer, cooler, brighter and etc. Note that all of these are language based and can vary in a diverse way through different cultures. Therefore, in order to describe colour in a systematic and

[1] A blind person was asked about what he felt of scarlet red; "Sound of trumpet" he answered.

mathematical way, accurate perceptual definitions are provided which are universally agreed upon . These definitions are in correlation with the physical measurable quantities. So far, precise definitions of colour are developed by the International Lighting Vocabulary, the ASTM standard [6] and most importantly of all the Commission Internationale de l'Eclairage (CIE) system [7].

Mixing coloured dyes or paints is an example of *subtractive mixing* while *additive mixing* is the combination of coloured lights. For additive colour mixing, colour stimuli can be matched by the additive mixture of three appropriately chosen primaries. In the CIE system this notion is used as a method of specification; that is, the amounts of the three primaries that are used to match a stimulus can form a specification of that stimulus.

2.2.1 Colour attributes

Colour is defined as the visual perception attribute which consists of combinations of chromatic and achromatic contents which can be described by names such as red, green , blue, white, pink, black, brown etc. Further specification of the quality of the colour is made by the words such as light, dark, bright etc. This definition of colour is accompanied by the critical note that the perceived colour depends on many factors, such as the size, shape, structure, surround, observer's experience, visual system adaptation etc. Hue is an attribute of a visual sensation and defines if an area is similar to one of the perceived colours: red, green, yellow, blue an etc. or a combination of them. According to this, colours can be categorised as chromatic or achromatic simply based on the existence of hue. Brightness is another attribute of colour which is concerned about the amount of light emitted from that area. Lightness is the amount of brightness of a stimuli compared to the brightness of a similarly illuminated area which appears white. So the difference between brightness and lightness is that brightness is used for absolute perception while lightness can be relative brightness.

However, in colour perception it is broadly accepted that three terms are needed to define colour. And therefore three of the colour attributes are usually combined to describe colours. However, some believe that five perceptual dimensions are required for a complete specification of colour appearance; brightness, lightness, colourfulness, chroma and hue. The selection of the attributes depends on the application and the information required.

2.2.2 Colour order systems

Colour order systems provide efficient data and techniques for colour appearance specification. In the colorant mixing system the main focus is on the relationships between colour primaries[2] and their mixtures [8]. Most colour wheels[3], for example, are based on this principle. In printing, *pantone* is a well known colour naming system (it has a greater colour gamut[4] than CMYK) and is based upon colour mixing. By contrast, the Munsell Color System (1905), which was one of the first colour-order systems[5], is developed based on colour perception. In Munsell system colour is specified according to the three attributes: hue, value and chroma. The reason why the Munsell book of colour became very successful was that it arranged colours with equal visual increments along the dimensions. Therefore, it indicated colours both psychophysically and physically. Munsell value, for example, is arranged in 10 main steps from black to white. Purple, blue, green yellow and red constitute the five main steps of the Munsell hue, plus the intermediate hues which are placed in between two hue steps (resulting in ten steps for Munsell hue). Having said that the colour samples in the Munsell system are evenly spaced in perceptual terms; Munsell Chroma is not an exception but there is no specific extreme point as the chroma depends upon the hue and value of that sample. However; being a cylindrical system, it has a limitation of non-uniformity[6] [9].

Other colour systems also provide visual colour stimuli such as the Swedish Natural Colour (NCS system), colour curve, OSA, DIN and Ostwald and have been developed and applied to different experiments, colour communications, model testing etc. Although, these systems make colour communication easier the viewing condition need to be considered as it affects the colour appearance. Many variables are involved in how the colours appear and small changes in each can result huge differences in colour appearance which are pointed

[2] The set of colorants used for a particular coloration process is referred as primary set; for example, cyan, magenta, yellow and black are primaries in printing and red, green blue are used in cathode ray tubes.
[3] Colour wheel is a circular presentation of the colour spectrum mainly used by artists and designers. Colour wheels can display from three or more colours.
[4] Colour gamut is the range of colours which a coloration system is produced in.
[5] Colour-order system is a conceptual system of organized colour perception.
[6] Comparing to CIELAB which is an approximation of a visually uniform colour-order system based on rectangular dimensions of lightness, redness-greenness and yellowness-blueness which will be described further on in 2.3.2.5.1.

2.2.3. That is why colour appearance models (discussed in 2.3.2.5) are proved to communicate colours more precisely.

2.2.3 Colour appearance elements and the visual field

The viewing field of an experiment consists of stimulus, proximal field, background and the surround. In general (and so it is here), the stimulus refers to a patch, coloured uniformly and about 2° angular subtense [8]. It is important to note that the appearance attributes of stimuli not only depend on its own specific attributes but also of the surround and other stimuli which are set nearby. Therefore, the placement of the stimulus in the visual field is very important. However, real life scenes are observed differently. For example; when looking at an automobile, the observer would assign a single colour to it by saying it is blue, although different parts of the car would have significant difference in colour attributes when viewed in isolation. So although the car is not viewed in a 2° angular field, it is viewed as a whole rather than a single stimulus. Apparently, many car companies use a single colour patch to illustrate the different available colours (specifically for online customers)[7]. Therefore, single colour patch presentation is the typical demonstrative approach. It is important for the experiments carried out in this thesis to use the same single coloured patches as colour stimuli. The nature of the online presentation requires these stimuli to be displayed on the monitor (in RGB mode) rather than physical samples. Because of this all the experiments involve digital colour stimulus that are displayed on the monitor in isolation. Proximal field is defined as the surroundings of the stimulus immediately observed after the stimulus. The Background is the actual environment where the stimulus is presented and is extended in all directions. A more precise definition is the 10° field starting from the edge of the colour stimuli.

The surround of the experiment is the area beyond the background where the stimulus is presented; it could be the room or laboratory. Specifying the surround plays an important role in modelling long-range induction, flare (stimulus and within the eye) and overall image contrast effects [10]. For most colour appearance models, absolute luminance or illuminance[8] units of the surround are required. CIE has defined a surround ratio S_R, which

[7] This is not limited to car sales only, since most of the products that have an on-line market are illustrated in the same way.

[8] Luminance is the colour perceived as belonging to a source of light whilst Illumination is the light source used for perceiving a colour.

is the ratio of the luminance of a white object in the surround area by a reference white on the display itself. The numerical value of the S_R is used to categorise average, dim and dark surround[9]. By controlling the lighting conditions of the surround of the stimulus more attention and focus is driven to the stimulus. Also the colour of the stimulus is less interrupted by nearby coloured objects that are luminated by the existing light of the surround. Although, most colour appearance experiments involve precise considerations of the surround (i.e. making the surround very dark), the real world visual field is not similar. It is very rare that people would view colours in a very dark room. Therefore, consideration of the uncontrolled surround, in which all possible viewing situations may occur, would be useful. However, the correlation between model predictions and visual evaluations is influenced by the visual geometry of the colorimetric measurements which varies when the experiment is run in uncontrolled conditions. Therefore, other than the elements described above, the adapting stimulus is required which in some models it is considered to be the background and in others it is the measure of the light source[10]. In other words it is an absolute measure of the tristimulus[11] values of the luminance or a white object under the luminance. So to summarise, depending on the colour model used; it is necessary to have information about the tristimulus values of the stimulus, proximal field, background and adapting field (measured according to the actual light source). Whilst the colour of the surround is not important, the relative surround luminance is required which can be approximated to 0% of the luminance of the white when the surround is so-called dark, 0%-20% when it is dim and 20%-100% when it is average. The importance of this specification is due to the light adaptation[12]. The chromatic adaptation is another element which should be considered as it is the independent changes of the responsivity amongst the three types of cone photoreceptors which includes the 'Afterimages' phenomena. That is when, for example, we stare at a red colour patch, and the eyes get adapted as the retinal areas exposed to the red become less sensitive. Therefore when moving on to a white area,

---

[9] Average surround has an S_R between [0.2,1), dim is (0,0.2) and dark is almost 0. Additionally a bright surround is defined for S_R greater than 1.

[10] The change in sensitivity to variation of colour attributes of a stimulus is called the adaptation. There are three types of adaptation; light, dark and chromatic.

[11] Three values used to describe colour.

[12] Light adaptation refers to the situation when increase in the overall level of illumination results in visual sensitivity decrease. Dark adaptation is the same but with the inverse direction. Which is thought to happen because of change in pupil size.

shades of cyan are visualised due to lack of equal sensitivity in the retina. This is why experiments that involve a sequential display of the colour stimuli, have a pause between the displayed colours. This ensures that if there is an after-image effect for any observer, it would not influence the appearance of the next colour directly as it will eventually disappear during the pause.

2.3 Fundamentals of Colour Science

Colours are named and communicated differently around the world. Some believe that there are an infinite number of different shades for the visible colours, which in the first instance makes colour identification impossible if it is only limited to visual naming. Therefore, colours are categorised according to their application. Primary colours for example are colours that their mixture results in a certain range of colours[11]. For example, artists (painters to be more specific) tend to categorise colours in three main primaries of red, yellow and blue, which are used for dyes and paint. Whilst colour scientists define two different colour mixing systems; the additive system in where as overlapping coloured lights (red, green and blue primaries) are mixed and the subtractive system, where mixtures of dyes and pigments (cyan, magenta, yellow and black primaries) are used (e.g. in printing) [11]. Colours can be described effectively and in a systematic way throughout the world by being coded as numbers in a three-dimensional system. Different standards have been set to define the three dimensions of colour with specific values; note that, standards are selected according to the application of colour measurements. In this section, the fundamentals of colour and colorimetry are presented along with a brief overview about colour appearance models.

2.3.1 Colour physics

Modern colour physics originated when Newton studied the colour spectrum[15] by reflecting a ray of white light using a prism in 1730 [8]. Light has a finite speed of $2.997\times10^{8}ms^{-1}$ in empty space but the energy of light varies as described by wave theory; “wave theory describes radiation in terms of electric and magnetic fields which resonate at right angles to each other and to the direction of motion” [12]. Electromagnetic waves are differentiated by their length which is the distance between two peaks in the wave; some reach up to

[15] Note that Newton was not the first to observe the spectrum; however, he drew some original and startling conclusions from his observations.

kilometres and metres (e.g. a metre long wave is called radio wave), whereas cosmic rays have the shortest wavelengths ($10^{-5} nm$). Visible spectrum is observed approximately between 380 and 780 nanometres [13].

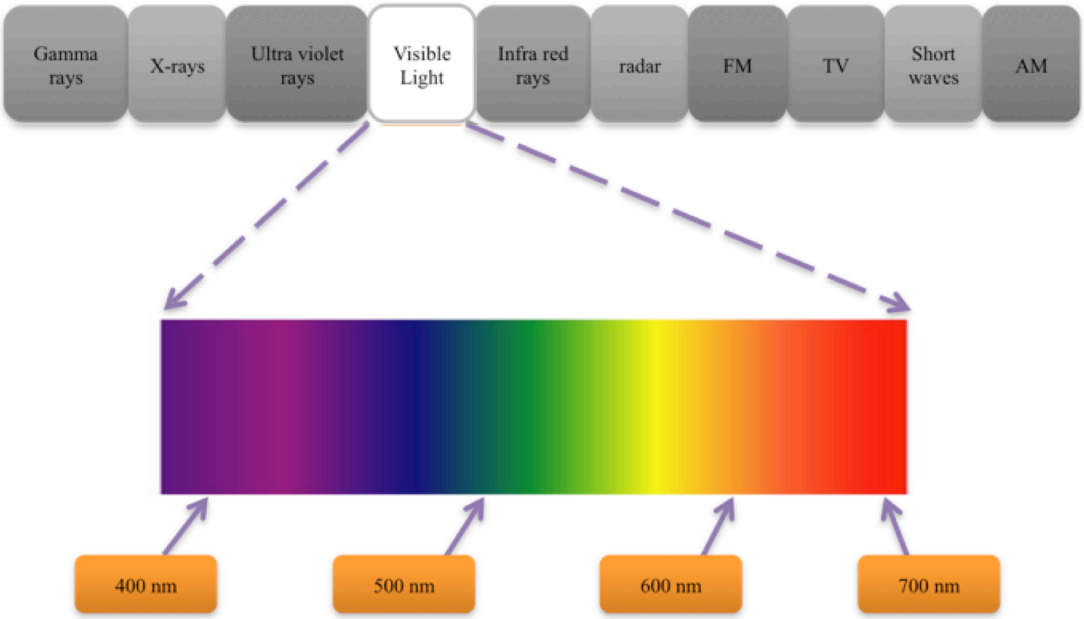


Figure 1:Location of the visible spectrum amongst electromagnetic radiation waves.

A surface that absorbs certain wavelengths of light striking it and reflects a portion of it appears coloured; a black surface absorbs all the light whereas a white surface reflects almost all of the light back [14].

2.3.2 Colorimetry

Colour appearance requires fundamental knowledge of Colorimetry, which by definition, means colour measurement, and its ultimate goal is to assess the appearance of colour stimuli by taking all sorts of real life variables (such as the surroundings) into account. The CIE system (International Commission on Illumination) which was first established in 1931, is the well known colorimetry standard. This system is used as the starting point for colour appearance models.

Colour is a visual sensation and its appearance greatly depends on three components. The light source, the eye and the coloured object. The quantification of these three components and their relationships makes the science of Colorimetry, which involves a wide range of fields such as physics, chemistry, psychophysics, physiology and psychology.

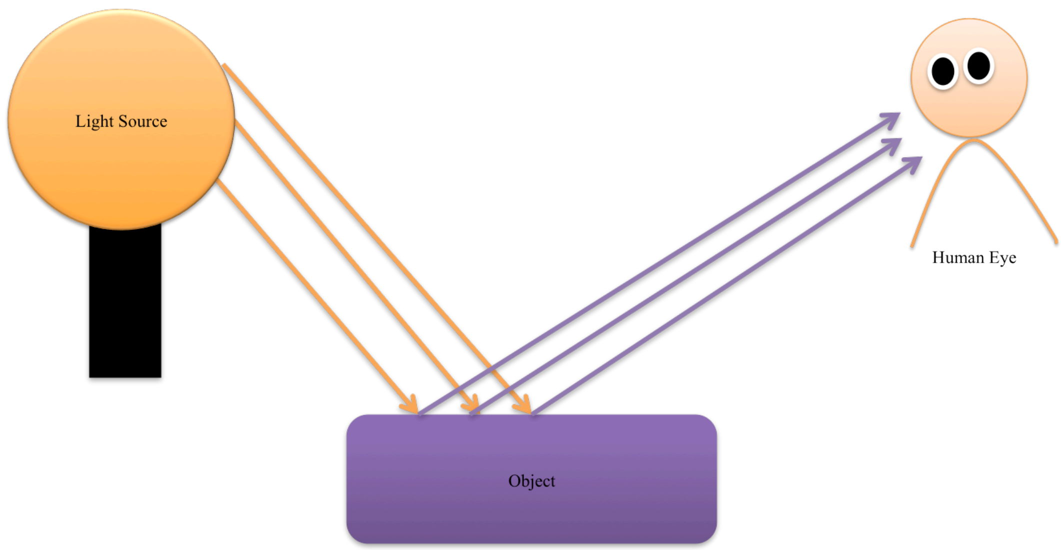


Figure 2: The three components of colour vision (known as the triangle of colour).

The light source and object properties would be explained in 2.3.2.1 followed by the deeper explanation of the CIE system. Finally, colour appearance and difference models are explained.

2.3.2.1 Light sources and illuminants

The very first feature of the triangle of colour provides the energy required for vision as it releases electromagnetic waves which hit the object; once the waves are reflected from the object the eye receives and converts these waves into images on the retina. In other words, the light source is a physical emitter of visible energy such as Incandescent light bulb, the sky or fluorescent tubes. Specific standards have been set by the CIE system, named as Illuminants, that some are measured values of the spectral power distribution[16] such as the CIE illuminants A which represents typical incandescent. However, in some cases the opposite is true.

[16] Spectral power distribution is referred to the distribution of the measured power for an illuminant per wavelength. This is usually presented in a plot or table.

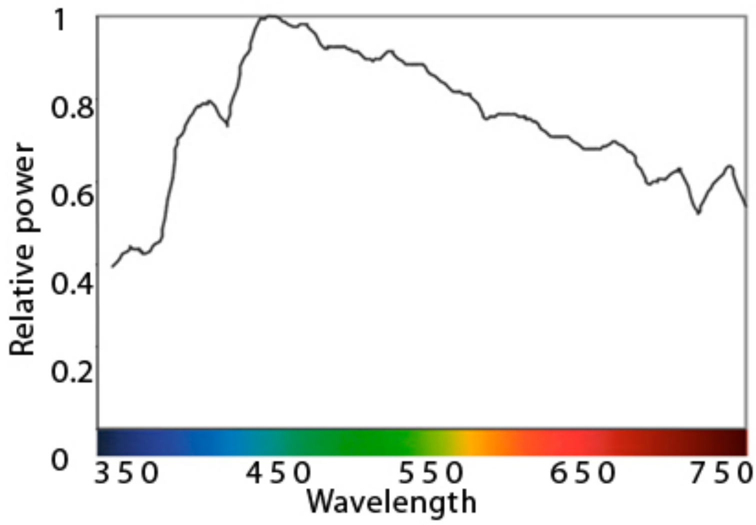


Figure 3: Spectral power distribution of CIE Illuminant D65 [15].

2.3.2.2 Coloured objects

The second feature of the triangle of colour is the object itself, as its properties influence the way it absorbs, reflects or transmits the radiant energy and therefore the way it is visualised differs accordingly. Reflection, absorption and transmission are defined respectively as the ratio of the energy reflected, absorbed and transmitted to the incident energy. Measurement of these three properties is the subject of spectrophotometry. It should be noted that reflection, absorption and transmission properties of objects not only depend on the wavelength but also the viewing geometry. This leads to an extreme complication of the spectrophotometry of glossy materials or even metallic or pearlescent finishes (used mainly in the car industry), as it requires vast measurements across all viewing angles as well as various wavelengths. Therefore, measurements on such cases are very expensive and difficult. It is also important to note that the surface of different objects reflect light differently at different angles. However, some differences exist between the simulating light source and its CIE illuminant spectral power distribution, which suggests that the actual spectral power distribution of the light source should be used. Measurements made on the power distributions of light source (per wavelength) are called spectroradiometry. Indeed, depending upon the nature of the object there are two kinds of reflection, specular and diffuse; specular reflection is divided in to three categories, which are displayed in Figure 4.

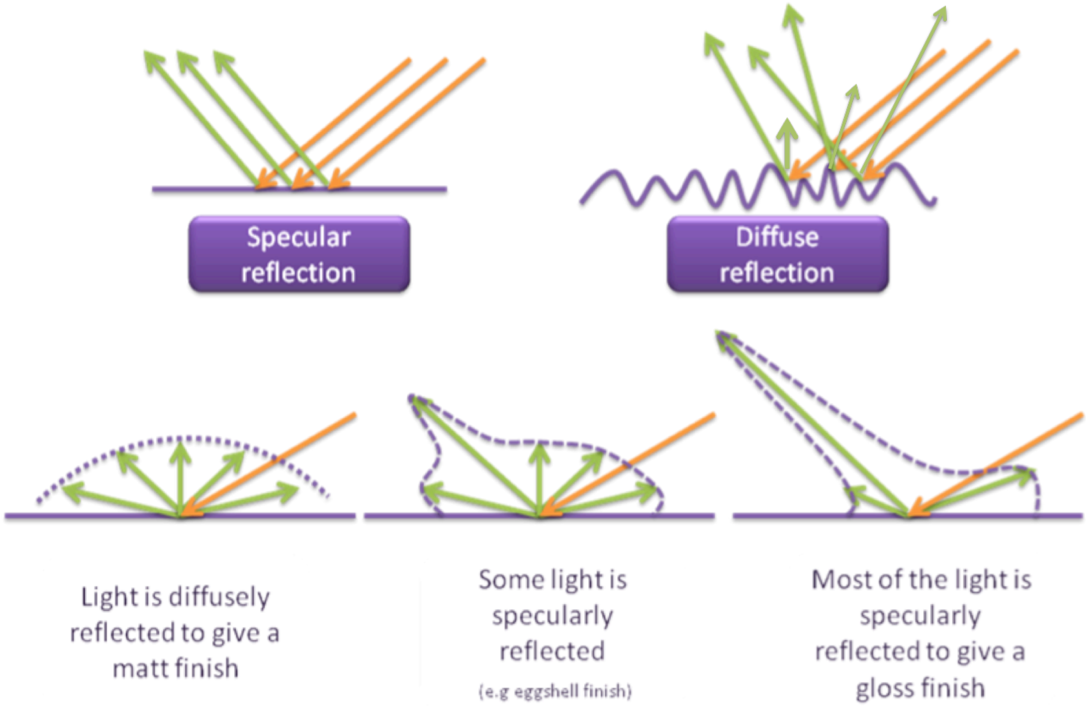


Figure 4: Specular and diffuse reflection

2.3.2.3 Visual response of the human eye

The visual response of the human eye is recognised by the colour matching phenomena. Since, the three cone types (long, medium and short wavelengths) will match two stimuli with different spectral power distributions of $\varphi_1(\lambda)$ and $\varphi_2(\lambda)$, only if the following assumptions hold[17].

Equation 1

$$\int \varphi_1(\lambda)L(\lambda)d\lambda = \int \varphi_2(\lambda)L(\lambda)d\lambda$$

Equation 2

$$\int \varphi_1(\lambda)M(\lambda)d\lambda = \int \varphi_2(\lambda)M(\lambda)d\lambda$$

Equation 3

$$\int \varphi_1(\lambda)S(\lambda)d\lambda = \int \varphi_2(\lambda)S(\lambda)d\lambda$$

[17] This is also known as the definition of Metamerism.

Where λ is the wavelength and L, M and S stand for long, medium and short wave length cone receptors respectively. At first, a system of colorimetry was formed based upon the principles of trichromacy and Grassman's law. Any colour can be matched by a linear combination of three other primary colours[18][8]; provided that none of those three can be matched by a combination of the other two. This is fundamental to colorimetry and is Grassman's first law of colour mixture. So colour S can be matched by R units of red ($\mathbb{R}$), G units of green ($\mathbb{G}$) and B units of blue ($\mathbb{B}$): $\mathrm{S} \equiv \mathrm{R}[\mathbb{R}] + \mathrm{G}[\mathbb{G}] + \mathrm{B}[\mathbb{B}]$. A mixture of any two colours (sources S1 and S2) can be matched by linearly adding together the mixtures of any other three colours that individually match the two source colours. This is Grassman's second law of colour mixture. It can be extended to any number of source colours:

Equation 4

$$\mathrm{S3}[\mathbb{S}3] = \mathrm{S1}[\mathbb{S}1] + \mathrm{S2}[\mathbb{S}2] = (\mathrm{R1} + \mathrm{R2})[\mathbb{R}] + (\mathrm{G1} + \mathrm{G2})[\mathbb{G}] + (\mathrm{B1} + \mathrm{B2})[\mathbb{B}]$$

In the CIE system, the amounts of the three primaries that are used to match a stimulus can form a specification of that stimulus. The CIE primaries are known as X, Y and Z[19] and the XYZ values (the so-called tristimulus values) can actually be calculated if the reflectance factors $P(\lambda)$ of a colour sample are known. The formula for calculating CIE tristimulus values for a surface with spectral reflectance $P(\lambda)$ under an illuminant of relative spectral power $\varphi(\lambda)$ is

Equation 5

$$X = k \int \varphi(\lambda) P(\lambda) x(\lambda) d(\lambda)$$

Equation 6

$$Y = k \int \varphi(\lambda) P(\lambda) y(\lambda) d(\lambda)$$

[18] The set of colorants used for a particular coloration process is referred as primary set; for example, cyan, magenta, yellow and black are primaries in printing and red, green blue are used in cathode ray tubes.

[19] The primaries XYZ were initially identified after the primaries red, green and blue were proven to be inefficient since, they required; 1- Negative values for the colour matching functions. 2- To force one of the functions to be equal to the CIE 1924 photopic luminous efficiency function. The first was solved by setting the primaries to imaginary colours which were more saturated than monochromatic lights. For overcoming the second problem, Y was chosen to be the only primary that holds the luminance response.

Equation 7

$$Z = k\int \varphi(\lambda)P(\lambda)z(\lambda)\,d(\lambda)$$

where k equals to 683 $lumen/W$ for absolute colorimetry whilst it is $100/[\int y(\lambda)\varphi(\lambda)d(\lambda)]$ for relative colorimetry. In this formula, the colour-matching functions – $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ – represent the response of an average observer [11].

2.3.2.4 Chromaticity diagrams

Tristimulus values are defined as the triplets which specify colour of a stimulus. The triple nature of the tristimulus values introduces colour in a three-dimensional space which to some extent can be inconvenient. Therefore, chromaticity diagrams were developed to aid the visualisation of colours in a two-dimensional space by a normalisation transformation on the tristimulus values[20], resulting the chromaticity coordinates, xy, thus[21]:

Equation 8

$$\bar{x} = \frac{X}{X+Y+Z}$$

Equation 9

$$\bar{y} = \frac{Y}{X+Y+Z}$$

Plotting y against x yields a map of colour space called the chromaticity diagram (Figure 5) which is a two-dimensional diagram with constant luminance. This diagram makes Grassman's law explicit, since the additive mixtures of two colours fall on a straight line joining the points that represent two colours.

The chromaticity diagram has certain limitations. For example, it is two-dimensional whereas all inclusive colour spaces are three-dimensional. Also it is visually, (perceptually), non-uniform since small changes in the coordinates can result dramatic differences between appearance of colours.

---

[20] This transformation eliminates information relevant to luminance, and therefore; colours from the three-dimensional space are all projected on a specific surface which is presented in a two-dimensional space.

[21] The third coordinate can be calculated by $\bar{z} = 1 - X - Y$.

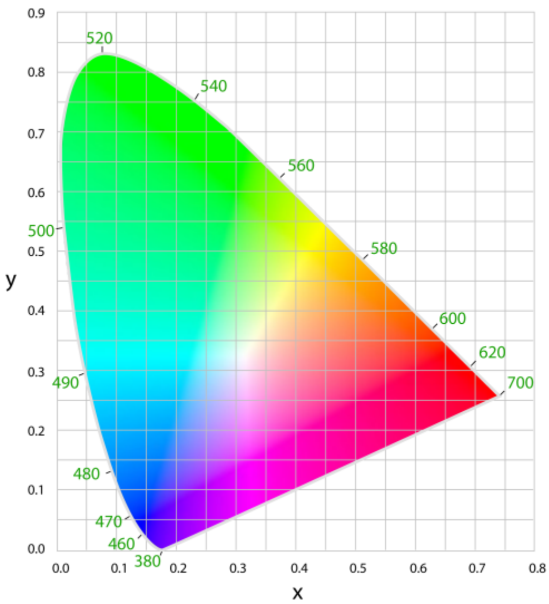


Figure 5: The Chromaticity diagram.

2.3.2.5 Colour appearance models

There are so many attributes to colour which need to be organised in order to make a well-connected description of colour, perceptually and numerically so that it can be communicated more easily and effectively. Various colour spaces have been developed such as CIELAB and CIELUV and finally CIECAM02 which will be described in the forthcoming section. A colour appearance models' formation is based upon a few concepts. All start with taking the tristimulus values of the stimulus and white object as inputs (XYZ and $X_nY_nZ_n$)[22]. Additional inputs such as the colorimetric measurements of the proximal field, background, surround luminance and other spatial or temporal information are required. Then a linear transformation of the tristimulus values takes place to result the cone response, precisely modelling the physiological procedure of the human vision. A Chromatic adaptation transform is the next important step leading to the predictors of colour appearance attributes.

CIELAB was the first uniform colour space to be developed; however, its limitations encouraged the development of more flexible and precise colour appearance model such

[22] When using colour measurement devices such as the spectroradiometer, tristimulus values are resulted which are the main reason of counting them as starting points.

as CIECAM97s and initially CIECAM02. CIECAM97s was a very successful colour appearance model developed by the huge demand made in the product design industry in 1996; identified by Hunt [16]. CIECAM97s was heavily applied by scientists and the industry and therefore, its limitations were identified rapidly. This lead to the development of CIECAM02 [17, 18]. CIECAM02 was simpler than CIECAM97s and has been shown to be the best model.

2.3.2.5.1 CIELAB colour space

Three-dimensional colour spaces have been developed by the CIE where the relevant dimensions are in approximate correlation with perceived lightness, chroma and hue . Therefore, a more uniform colour space is provided; CIELAB space[23][10] is developed based on a transformation of the tristimulus values (X, Y and Z) which are normalised according to the tristimulus values of the reference white point (X_n, Y_n and Z_n)[24]. The following equations[25] are used for deriving the L^*, a^* and b^* values [11].

Equation 10

$$L^* = 116(Y/Y_n)^{\frac{1}{3}} - 16$$

where $\left(\frac{\mathrm{Y}}{\mathrm{Yn}}\right) > 0.008856$ and

Equation 11

$$L^* = 903.3(Y/Y_n)^{\frac{1}{3}}$$

where $\left(\frac{\mathrm{Y}}{\mathrm{Yn}}\right) \ll 0.008856$

Equation 12

$$a^* = 500\left[(X/X_n)^{\frac{1}{3}} - (Y/Y_n)^{\frac{1}{3}}\right]$$

where

[23] Note that CIELAB was developed specifically for colour difference between two stimuli rather than colour appearance space, but since it provides colour appearance attributes of each of the two stimuli, with caution; it can be named as a colour appearance space.

[24] I.e., X/X_n, Y/Y_n and Z/Z_n.

[25] Note that these are not the full equations.

Equation 13

$$b^* = 200\left[(Y/Y_n)^{\frac{1}{3}} - (Z/Z_n)^{\frac{1}{3}}\right]$$

in which L^*represents lightness, a^* represents redness – greenness and b^* represents yellowness – blueness. These coordinates develop a three-dimensional Cartesian space for colour, which is illustrated in Figure 6.

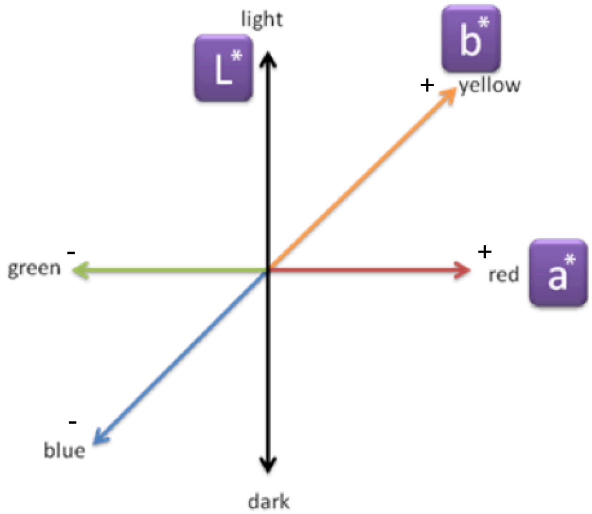


Figure 6: Representation of the L^*, a^* and b^* coordinates.

CIELAB and CIELUV[26] colour spaces are only approximately uniform and can be inadequate for specific applications [19], and its non-uniformity applies to different regions of the colour space [20-22]. Other limitations of the CIELAB 1964 space in brief are the wrong von Kries transformation [23], the un-alignment of the perceptual unique hues (red, green, yellow and blue) with the a^* and b^* axes. Also straight lines for the constant perceived hue appear as curved in the CIELAB space (see [24] for more details). Therefore, other systems have been suggested by the CIE which are based on the CIEAB space (since the CIELAB has been accepted and applied widely) but with a Cartesian approach. For example the LCH

[26] CIELUV was introduced for self-luminous colour stimuli which are generated by additive colour mixing:

$L^* = 116(Y/Y_n)^{1/3} - 16$
$u^* = 13L^*(u' - u'_{\mathrm{n}})$
$v^* = 13L^*(v' - v'_n)$

space which is a non-Euclidian conversion of the CIELAB space, specifying colour in terms of lightness, hue and chroma.

Equation 14

$$L^* = 116(Y/Y_n)^{1/3} - 16$$

Equation 15

$$C^*_{ab} = \sqrt{\left({a^*}^2 + {b^*}^2\right)}$$

Equation 16

$$h_{ab} = \tan^{-1}(b^*/a^*)$$

The L^*, C^*_{ab} and h_{ab} coordinates are used for a cylindrical representation of the CIELAB space.

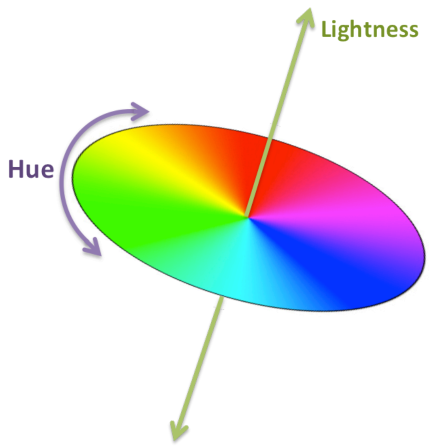


Figure 7: Presentation of lightness and hue.

CIELAB should only used for comparisons between colour stimuli (of same size and shape) which are viewed under the same conditions of surround luminance and background colour. The observer should be adapted to a field of chromaticity not very different from average daylight. However, CIELAB is not able to encounter the spatial variables which effect the colour appearance. For example; an orange colour patch which is located on two different backgrounds (as in Figure 8) would have the same measured CIELAB values whilst it appears differently on each background.

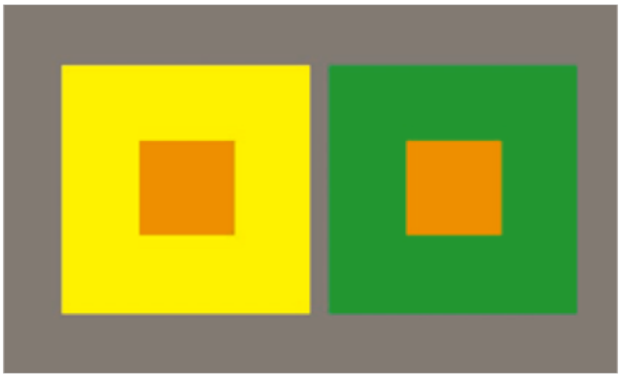


Figure 8: The orange colour patch looks darker on the yellow background compared to the green background.

This goes back to the concept of relative colours, when the colours are not viewed in isolation. In this case additional variables, which hold the background of the patch, and also other spatial variables should be involved. All of these limitations encouraged the development of more flexible and precise colour appearance models such as CIECAM97s and initially CIECAM02.

2.3.2.5.2 CIECAM02

The huge success of CIECAM97s encouraged more improvements in the model, leading to the development of the CIECM02. The input data of this model are the relative tristimulus values (XYZ) and the white point ($X_w Y_w\ Z_w$), the adapting luminance $\mathrm{L_A}$ [27], the relative luminance of the surround (dark, dim or average[28]) and finally a choice whether a discount applies to the illuminant or not[29]. According to the identified surround relative luminance, three essential variables (c, N_c, F) are revealed.

[27] The adapting luminance is usually set to 20 percent of the luminance of a white stimulus present in the experimental scene.
[28] Average corresponds to reflection prints, dim to television or CRT displays and dark to projected transparencies.
[29] It is discounted mainly in cases were object colours are of interest and not for displays.

Table 1: CIECAM02 input variables.

Viewing condition	c	N_c	F
Average surround	0.69	1	1
Dim surround	0.59	0.9	0.9
Dark surround	0.525	0.8	0.8

A linear Von Kries chromatic adaptation transform has been used which is the main cause of its simplicity [25]. In addition it allows model inversion. Hue angle h is measured in the very same way as the in CIELAB and is ranged between 0° to 360°. Hue quadrature *h* and eccentricity factor *e* are also calculated using the following equations.

Equation 17

$$h = \tan^{-1}(b/a)$$

Equation 18

$$e_t = 1/4[\cos\left(H\frac{\pi}{180} + 2\right) + 3.8]$$

Equation 19

$$H = H_i + \frac{100(h - h_i)/e_i}{\frac{h - h_i}{e_i} + \frac{h_{i+1} - h_i}{e_{i+1}}}$$

Achromatic response is calculated using the following;

Equation 20

$$A = [2R'_a + G'_a + \left(\frac{1}{20}\right)B'_a - 0.305]N_{bb}$$

Lightness J is derived using the achromatic signals f the white point and stimulus;

Equation 21

$$J = 100(A/A_W)^{cz}$$

ϱ stands for brightness, which is calculated from lightness and the achromatic for white.

Equation 22

$$\varrho = \left(\frac{4}{c}\right)\sqrt{\frac{J}{100}}(A_W + 4)F_L^{0.25}$$

t is a temporary quantity, that relates to saturation which takes the surround and background's chromatic induction factors ($N_c$andN_{cb}) in to account;

Equation 23

$$t = \frac{(50000/13)N_c N_{cb} e_t \sqrt{a^2 + b^2}}{R'_a + G'_a + \left(\frac{21}{20}\right) B'_a}$$

Chroma *C*, colourfulness *M* and saturation *s* are derived from the following equations.

Equation 24

$$C = t^{0.9} \sqrt{\frac{j}{100}} (1.64 - 0.29^n)^{0.73}$$

Equation 25

$$M = CF_L^{0.25}$$

Equation 26

$$s = 100\sqrt{M/\varrho}$$

The equivalent Cartesian coordinates of chroma, colourfulness and saturation are $a_c = C\cos(h)$, $b_c = C\sin(h)$, $a_M = M\cos(h)$, $b_M = M\sin(h)$, $a_s = s\cos(h)$ and $b_s = s\sin(h)$. Note that the simplistic transformation of the chromatic adaptation applied for the CIECAM02 has caused a significant improvement in the inversion of the colour appearance model, which is useful in colour reproduction. The CIECAM02 can predict most of the phenomena such as the correlation of lightness, brightness, chroma, colourfulness, saturation and hue. It can also predict the effects of factors dependent on adaptation, surround and luminance. However it is not valid for calculations on very high or low luminance levels.

2.3.2.6 Colour differences

Euclidean differences between colours are measured using the following equation in CIELAB space;

Equation 27

$$\Delta E^*_{ab} = \sqrt{\left(\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}\right)}$$

Although the CIELAB space was developed in order to approximate a uniform space, it appears that the colour differences derived from the ΔE^*_{ab} formulae are not uniform. Modified versions of the ΔE are based on the computational conversion of the formulae in a Cartesian manner so the difference between the lightness, hue and chroma values are calculated with different weights forming several colour difference formulae such as CMC (l:c) distance function [26], the BFD (l:c) function [27, 28] and finally the CIE94 colour difference formula [29] displayed below;

Equation 28

$$\Delta E_{94}^{*} = \sqrt{(\frac{\Delta L^{*}}{K_L S_L})^2 + (\frac{\Delta C_{ab}^{*}}{K_C S_C})^2 + (\frac{\Delta H_{ab}^{*}}{K_H S_H})^2}$$

Equation 29

$$\Delta H_{ab}{}^{*} = \sqrt{\Delta E_{ab}^{*}{}^{2} - \Delta L^{*2} - \Delta C_{ab}^{*}{}^{2}}$$ [31]

Equation 30

$$S_L = 1$$

Equation 31

$$S_C = 1 + 0.045 C_{ab}^{*}$$

Equation 32

$$S_H = 1 + 0.015 C_{ab}^{*}$$

The variables K_L, K_C and K_H adjust the weight for lightness, chroma and hue. Interestingly the average value of ΔE_{94}^{*} tends to be less than $\Delta E_{ab}^{*}$for the very same colour pairs. However, there are several limitations caused by the specification of the reference conditions of use. The recently developed CIEDE2000 colour difference formula which performs better than CIE94 on small colour differences[32] [17] is an improved version of the colour difference equation [16, 18]. CIEDE2000 is used as a powerful industrial colour difference equation, which predicts colour differences between large colour patches very accurately and therefore is also used for difference images [21]. The main differences between the CIEDE2000 and the CIE94 are the adjustments made mainly for corrections on the hue differences. Also additional changes on the hue, lightness and chroma weighting functions have been made in order to improve the non linearity of the blue region [11, 23]. In this thesis the CIEDE2000 has been applied using the MATLAB toolbox [11].

---

[31] $\Delta H_{ab}{}^{*}$ is the difference between two hue angles. An alternative formulae is$\Delta H_{ab}{}^{*} = \sqrt{\Delta a^{*2} + \Delta b^{*2} - \Delta C_{ab}^{*}{}^{2}}$

[32] For colour differences larger than 5 the CIE recommends use of the ΔE_{ab}^{*}.

Equation 33

$$\Delta E_{00} = \sqrt{(\frac{\Delta L'}{K_L S_L})^2 + (\frac{\Delta C'}{K_C S_C})^2 + (\frac{\Delta H'}{K_H S_H})^2 + R_T(\frac{\Delta C'}{K_C S_C})(\frac{\Delta H'}{K_H S_H})}$$ [33]

Where

Equation 34

$$S_L = 1 + \frac{0.015(L'-50)^2}{\sqrt{20+(L'-50)^2}}$$

Equation 35

$$S_C = 1 + 0.045C'$$

Equation 36

$$S_H = 1 + 0.015C'T$$

Equation 37

$$\Delta L' = L'_T - L'_S$$ [34]

Equation 38

$$\Delta C' = C'_T - C'_S$$

Equation 39

$$\Delta H' = 2\sqrt{C'_T C'_S}\sin(\frac{\Delta h'}{2})$$

Equation 40

$$\Delta h' = h'_T - h'_S$$

It is also important to note that in CIEDE-2000 notation we have $L' = L^*$, $a' = (1 + G)a^*$, $b' = b^*$, $C' = \sqrt{(a'^2 + b'^2)}$ followed by

Equation 41

$$h' = \tan^{-1}(\frac{b'}{a'})$$

Equation 42

$$G = 0.5 - 0.5\sqrt{\frac{C_{ab}^{*7}}{C_{ab}^{*7}+25^7}}$$

Equation 43

$$T = 1 - 0.17\cos(h' - 30) + 0.24\cos(2h') + 0.32\cos(3h' + 6) - 0.2\cos(4h' - 63)$$

R_T stands for the rotation which is calculated by

Equation 44

$$R_T = -2\sin(2\Delta\theta)R_C$$

[33] The calculations of CIEDE-2000 is carried out by the *cie00de* function in MATLAB.

[34] The S and T as subscripts in the formulae refer to standard and trial.

Equation 45

$$R_C = 2\sqrt{\frac{c'^7}{c'^7 + 25^7}}$$

Equation 46

$$\Delta\theta = 30\exp\left(-\left(\frac{h'-275}{25}\right)^2\right)$$

However, in some cases colour difference formulae may not be the only way to understand the difference between colours. Since, these formulae are based upon some optimisations of the colour space assuming the uniformity of the relative space. In other words, the precision of these formulae is related to the uniformity of the colour space. Therefore, an alternative method for defining colour difference (on the chromaticity diagram) is the graphical approach of MacAdam's ellipsoids which is based on the concept of colour specification. The subject of colour specification reveals whether two colours with different spectral distributions would look similar to the human eye or not. This eventually leads to the formation of ellipsoids that define the region in which all the colours are indistinguishable from the one located in the centre of the ellipsoid. The colours located on the boundary of the ellipsoid would have *just noticeable difference* with the colour in the centre. The size and shape of the ellipse depends on the colour itself and has been a subject of many investigations [19-21].

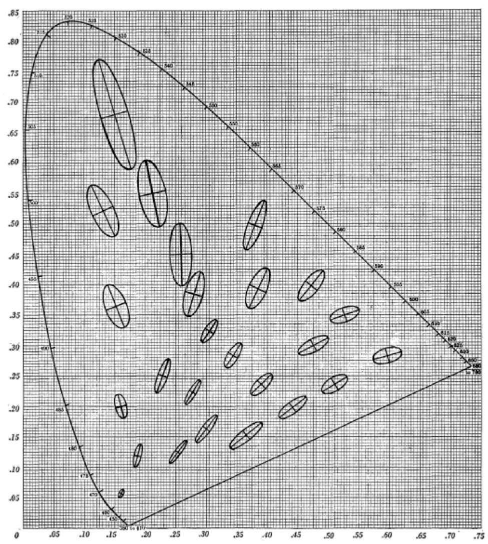


Figure 9: An example of the MacAdam's colour difference ellipsoids [21].

The application of the Mac Adam's visual sensitivity ellipsoids is proved to be convenient for cases in which large numbers of calculations of colour differences are required on relatively similar colours. The use of the colour difference formulae show more accuracy for cases with large colour differences (say the difference between red and orange) [22]. Moreover, the Mac Adam's ellipses are the cross sections of the three dimensional ellipsoids, which correspond to even luminance levels [26]. Therefore there is a lack of information consistency with this method. Some effort has been made to overcome this problem, by also considering the ellipse on the YX plane [22, 27]. The results for both graphical and formulisation of the colour difference correlate well with each other and therefore it is wise to choose the one which fits best to the criteria of the problem.

2.3.2.7 Display Characterisation

Colorimetry involves accurate measurement and calculations. Not only the visual field variables, but also the display conditions play an important role. Therefore, in colour experiments where any sort of imaging device (such as CRT, monitors, printers or cameras) is used, device calibration is applied. It is the act of setting the device into specific display states and ensuring that it would stay constant through time of use. The actual relationship between the coordinates measured from the device such as the RGB or CMYK values with the CIE XYZ values (which are independent of the device) is called characterisation [10, 30]. Three different types of characterisation are applied. These are the use of physical models (such as the Kubelka – Munk model [28, 29] or the GOG model), look-up tables and numerical methods [31]. Colour devices are categorised into three main groups of input, display and output. Characterisation models are defined differently for each of these devices. For transferring images between devices ICC profiling is required. ICC stands for the International Colour Consortium which has focused on the problems of communicating colour across open systems. It aims to ensure that the colour going through an input device is mapped properly to the output device colour space[35]. In its early days ICC profiling was mainly done by accompanying the input image file with extra information of the profile which was later on improved to more complex methods[36] [32]. In general, the science of colour management provides the translation technology which aids the transformation of

---

[35] Colour space is referred to the models which help in defining the colours numerically. Such as the RGB colour space. Since, colour is used across different devices, it needs to be defined with coordinates independent of the device.

[36] It caused increase in the size of the file; also many users did not require this level of flexibility and control. On the other hand, most existing file formats did not support colour profile embedding.

colours from a colour space of one structure (such as a coloured object, image, graphics or texts) to a colour space of an output (such as the monitor, printer, etc.). Therefore, ICC profiling as a field of colour management plays an important role in defining each device's standard systems. On the other hand, the development of the web has caused even broader use of images and these images are viewed on various displays around the world with different colour spaces. Coloured display monitors work by pixels. However, it is often misunderstood that the pixels represent the colour when actually they are the indicators of the colour lookup tables. A colour lookup table can consist of 2 to 256 colours which each of them are specified with red, green and blue components[37] [33]. In other words, the display colour space is an RGB based system and varies according to the display settings. Therefore there has been a need to merge the colour spaces of monitors and standardise a certain RGB colour space[38]. This standard was named the sRGB colour space which aimed to develop the colour fidelity of the desktop environment[39][32]. The development of the sRGB space is based on the mathematical transformations of the CIE colorimetric values, which is based on the averaged performance of all the computer displays and is calculated using the following;

Equation 47

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Applying sRGB transformations, ICC profiles, monitor calibration and in general colour management is required in order to achieve accurate colour management. Nowadays, web not only plays an important role in transferring knowledge but also is used as a marketing tool and therefore attracts many customers to products and forms online purchase. Accurate illustration of the colour of the products can only be accomplished through colour management which has become feasible by the development of various models [2]. However, colour management systems have also executed algorithms for Gamut mapping. For an additive device (monitor display) the maximum chromaticities that can be displayed

---

[37] Each of these components may be 8 bit deep.

[38] RGB space is device dependent that is why colours are not replicated in a constant manner across different devices.

[39] The standard colour space "sRGB" was offered by Hewlett-Packard and Microsoft, within the Microsoft operating systems, HP products, the Internet, and all other interested vendors.

are denoted by the triangle (called the colour gamut) with the three primaries indicating the three points of the triangle (see Figure 10).

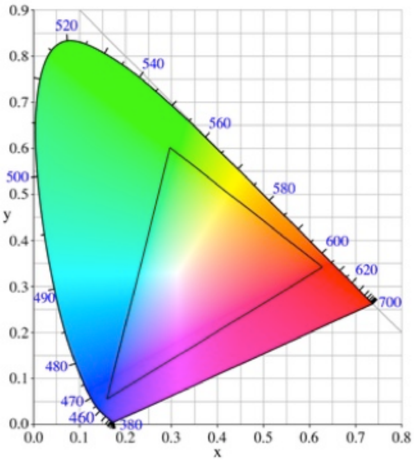


Figure 10: Colour gamut on the chromaticity diagram with red, blue and green as primaries.

Gamut mapping is the process of adjusting the colours according to algorithms which then allows original colours to 'fit' inside differently shaped colour gamut for successful transfer across media. For monitor displays; the colour gamut is set according to the factory default settings and therefore varies across different companies. This variety of setting, is one of the reasons why colour appearance is not equal on different displays [34].

3 Psychophysical and Statistical Methods

3.1 Psychophysical experiments and techniques

The science of psychophysics is about the relationship between a physical event and its corresponding psychological event in a quantitative manner [35], In other words it is the study of the relationships which exists between the perceptions evoked from a stimuli and its physical measurements [9]. In other words it is the study of the relationship between the stimulus (physical measurement) and response and also the theoretical notions about explanatory systems with the sensations and perceptions those stimuli evoke [9, 36]. Psychophysical experiments are designed to enable the psychophysical investigation. Throughout the 150-year-old history of psychophysics; Weber[40], Fechner[41] and Stevens[42] have made a huge impact in forming fundamental theories. Psychophysical experiments can be classified in to two broad categories: 1- Threshold and matching experiments. 2- Scaling experiments [9]. To be more precise, four different types of methods are used which are threshold, matching, measuring differences and direct ratio scaling. The first two, deal with intolerance of samples where the researcher investigates if there is specific difference between samples and standard stimuli; threshold technique finds the stimulus which is different from the standard whilst the matching method recovers the stimulus which can't be differentiated from a standard stimulus. Measuring difference technique is used when the strength of a specific feature is of interest, for example; when two colour stimuli are compared with each other on the basis of their hardness. The direct ratio scaling is used when the numerical values are assessed to the set of stimuli according to certain perceptual attributes, an example of this method is the Lutchi data set where observers were asked to scale the lightness, colourfulness and hue attributes for colours [35].

In this study, the difference measuring technique is used where observers are asked to rate each colour stimuli. The difference measuring technique can be performed in five different methods which are rank order, graphic rating, partition scaling, paired comparison and categorical-judgment. Rank order is used in experiments where observers are asked to

---

[40] Weber's law explains that the perceptual ratio $\frac{\Delta I}{I}$ remains constant (where I stands for the initial magnitude of the stimulus and ΔI is the change required to achieve a threshold).

[41] Fechner's law states that the logarithm of the physical stimulus intensity indicates the proportion of a stimulus' perceived magnitude.

[42] Steven's theory explains that the relationship follows a power law with different exponent for each perception.

arrange a given set of colours according to a certain attribute such as hue. Data gathered from this experiment are analysed to find ordinal or interval scales. Graphic rating technique is used when the experiment requires the observers to evaluate their perception on a one dimensional scale (e.g. a number from 0 to 10), whilst partition scaling is used when observers have to choose a stimulus which falls between two stimuli which are based upon a single attribute. These two scaling method are both used to provide a direct estimation of the interval scale. Comparative and categorical-judgment are explained in 3.2.1 and 3.2.2.

3.2 Scaling methods and theory

There are four main types of scales; 1- Nominal is used for categorical variables which are not ordered (for example, football players whom are numbered according to their place on the ground or colours which are named as red, green, yellow etc.). 2- Ordinal; a scale which only describes a specific order of greater or less (e.g. the arrangement of a specific hue according to the degree of lightness and darkness). 3- Interval scale which is a uniform space of differences but without a centre point such as zero (°F for example). 4- Ratio which is a uniform scale containing a centre point of zero (length for example) [35, 37]. In this study ordinal scaling is used and is formed as Likert scale which is a symmetric scale with equal number of positive and negative choices [38]. A typical Likert's scale is as followed

1- Strongly disagree
2- Disagree
3- Neither agree nor disagree
4- Agree
5- Strongly agree

Although some use 7 or 9 scales, empirical studies prove that the 5 point scales may perform better in terms of mean values [39], and therefore the above 5 scales are used in this study. Likert's scale takes a step forward compared to other comparative scales which are also bipolar scaling methods, measuring either positive or negative response to a statement. Often they are four-point scale and act as a "forced choice" method since the middle option of "Neither agree nor disagree" is not available [40]. But Likert scales may be subject to distortion from several causes. Respondents may keep away from extreme response categories which would cause central tendency bias, or agree with statements as presented (acquiescence bias); or try to portray themselves or their organisation in a more favourable light (social desirability bias). This problem can be prevented by designing a scale with balanced keying (an equal number of positive and negative statements), since acquiescence on positively keyed items will balance acquiescence on negatively keyed items, but central tendency and social desirability are somewhat more problematic. It is important to note that because the online survey is global and designed in different languages it is necessary to keep it as simple and least bias as possible. Since all the words in the survey were to be translated in the relevant language it would have been very complicated to have the 5 word formats of the Likert's scale as mentioned above so it was decided to use numbers which can be easily understood to all nations and communicates in a simple manner (See Figure 11).

Bi –polar is described by Oxford Dictionary as having two poles or extremities. The term bi-polar characteristic which is used throughout this text is used to note feelings or emotions which are arranged in a bi-polar manner. As an example, heavy-light can be a bi-polar characteristic of a product which is aroused from a person viewing that product in a certain colour. So for each of the bi-polar characteristics each person could rate their opinion on Likert's 5-point scale where 2 on the left indicates the highest degree of heaviness and 2 on the right indicates the highest level of lightness, the same is for the scale 1 and finally zero indicates the option of none of the bi-polar characteristics.

	2	1	0	1	2	
Heavy	○	○	○	○	○	Light
Warm	○	○	○	○	○	Cool
Modern	○	○	○	○	○	Classical
Clean	○	○	○	○	○	Dirty
Active	○	○	○	○	○	Passive
Hard	○	○	○	○	○	Soft
Tense	○	○	○	○	○	Relaxed
Fresh	○	○	○	○	○	Stale
Masculine	○	○	○	○	○	Feminine
Like	○	○	○	○	○	Dislike

Figure 11: The Likert scale used in this study.

In general, one dimensional-scaling methods are used in experiments when the relationships between perceptual magnitudes and physical measures of stimulus intensity are of subject. According to the aims and objectives of this research, category scaling and paired comparison are used. Category scaling is mainly used to obtain ordinal or interval scales when large number of stimuli are involved in the experiment. Typically, a categorical-judgmental experiment involves the categorisation of samples. Torgerson established the law of categorical-judgment in 1954; which assumed that the interval data derived from the categorical-judgment method have normal distribution [41]. Paired comparison technique is used when the samples are compared pair by pair, in which the proportion of times that a specific sample is judged higher (or lower) than the other sample in a specific attribute. Thurstone established the law of paired comparison in 1927 in order to obtain interval data from this experimental method [42].

3.2.1 Comparative-judgments

Comparative-judgment is mainly used when two sets of stimuli are compared with each other. It is also used for the comparison of bi-polar characteristics by Ou and et.al, when each of the observers had to choose between one of the bi-polar characteristics offered according to the presented colour (see 4.6.2). For example; the colour red is proposed and the observer has to decide if it is heavy or light so the main comparison is between the word pairs amongst the bi-polar characteristics. The data collected are ordinal and by using Thurstone's law of comparative-judgment they are converted in to interval scales [42], in other words the number of times that a certain word pair is chosen is converted to a portion which describes how often it has been preferred rather to the other. The use of this law on colour can be assumed as followed:

1- For each colour a process is formed by which observers make judgments on the bi-polar characteristics which is called a discriminal process.
2- The discriminal process can eventually form a normal distribution.
3- The average and standard deviation of this distribution are related to the value and discriminal dispersion of the stimuli.

Case V of the Thurstone's law, is expressed by the following formulae;

Equation 48

$$R_i - R_j = z_{ij}\sigma\sqrt{2}$$

The paired comparison method is well known as an accurate result provider whilst on the other hand it requires vast number of comparisons which makes the experiment longer [35]. A brief summary of the algorithm of the paired comparison method is; the matrix of the raw data are used in which the cells stand for the frequency of which a word pair (column) is preferred to another (row), this matrix is then converted to proportion matrix by division of the total number of observations and then are transformed by the use of the equation $LG_1 = \ln(\frac{f_{ij}+0.5}{N-f_{ij}+0.5})$, initially by plotting these values against the cumulative normal values of the proportion matrix a strong linear relationship should be identified. The slope value of the linear model is multiplied in to the former matrix (the LG matrix) to result the Z matrix which its values are called z scores and are used to calculate the mean and standard deviation values.

3.2.2 Categorical-judgments

In colour and vision research; when different stimuli perceptions are judged to fall into certain categories, Thurstone's law of categorical-judgment is used which is expressed in the following equation

Equation 49

$$B_k - R_j = z_{jk}({\sigma_j}^2 + {\sigma_k}^2 - 2r_{jk}\sigma_j\sigma_k)^{1/2}$$

Where B_k is the mean location of the κth category boundary; R_j is the mean response to stimulus j; $\sigma_k$is the discriminal dispersion of the κth category boundary; σ_j is the discriminal dispersion of stimulus j; r_{jk} is the coefficient of correlation between momentary positions of stimulus j and category boundary κ on the scale; z_{jk} is the normal deviate corresponding to the proportion of times stimulus j is placed below boundary κ. It is important to point out that the law of categorical-judgments are valid when with respect to the category boundaries, the relative situations of stimuli are of interest. The analysing process of the categorical methods defers from ordinary analysis in the sense that the percentage proportions have to be transformed into standard normal deviates (z scores). The algorithm of the categorical-judgment can be summarised as followed; the proportional cumulative frequency scores are calculated from the frequency values matrix, then they are converted in to Z scores by using the inverse equations of the standard normal cumulative distribution (via the $LG_1 = \ln(\frac{f_{ij}+0.5}{N-f_{ij}+0.5})$ formulae, in case of infinite values) which correspond to an interval scale. The matrix of differences is then calculated by subtracting two columns of the Z matrix, which corresponding boundary values are derived by calculating the mean values of each column and summing them up. The scale matrix is the product of subtracting the Z matrix from the boundaries. The mean calculated for each row of the scale matrix is ranked and can stand for the ranks of the stimuli see first Appendix. The categorical-judgment method can be described as an algorithm which takes the matrix of frequencies as the input and eventually returns the relative z scores (the algorithm has been programmed in MATLAB).

3.3 Statistical Methods of Analysis

Like all other scientific experiments, colour experiment methodologies involve statistical methods too. The nature of the data gathered from colour experiments defines the type of statistical methodology which should be used.

3.3.1 Factor analysis

The term semantic differential refers to meanings which are defined by a pair of antonyms that have 5, 7 or more possible ranks between them. The word pair is called semantic scale but in this thesis, it will be referred to as bi-polar characteristic since it will be used for describing the characteristics of colour [43]. Over hundreds of bi-polar characteristics have been studied up to now, which have made the research process lengthy. This is why categorising bi-polar characteristics in to groups and also reducing them in to smaller numbers would be useful. First correlating bi-polar characteristics can be identified.

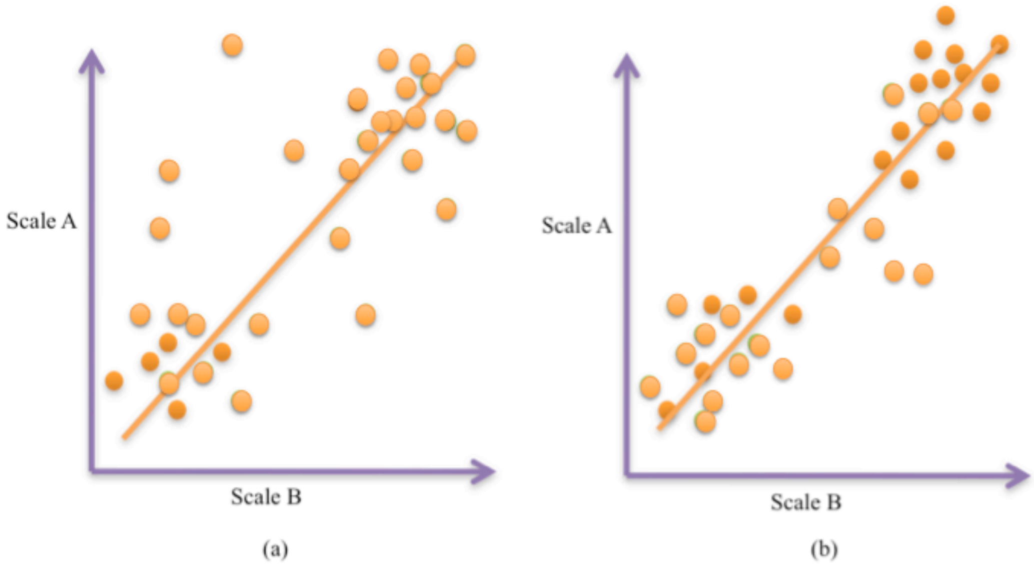


Figure 12: (a) Low correlation between two scales, (b) highly correlated scales.

However, highly correlated scales may not necessarily mean that the two scales are equivalent for all principles. Since correlation and causation are two separate concepts therefore, correlated scales may just act alike rather than affect each other. Also it is crucial to compare the scales with respect to their relative sensitivity in different parts of the range [43]. For example, in Figure 12(b) the correlation between the two scales is high but scale A gives higher accuracy in the higher levels while scale B has higher accuracy in the lower levels. Initially the significant components derived from factor analysis are used in modelling relationships.

Once correlation amongst the scales is identified, it will be possible to select a small number of scales that represent each other. By this the bi-polar characteristics which overlap each other in terms of meaning, are factorised in to a group named as components. More detailed discussion is made on this matter in 4.6.1 Bi-polar characteristics. Figure 13 (a) illustrates two components, which are not located in a parallel manner suggesting that the two components are unlikely to vary similarly. On the other hand, Figure 13 (b) illustrates two components, which are parallel suggesting that the two components follow a similar trend. This information, aids the interpretation of each component's attitude towards change.

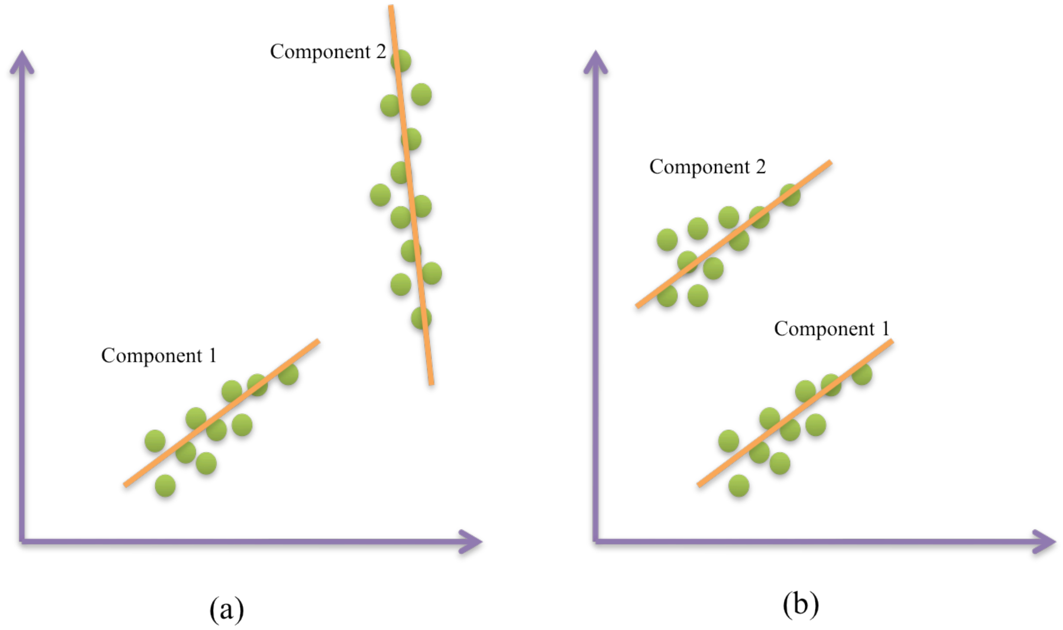


Figure 13: Factor plot for components 1 and 2: (a) Non-parallel components, (b) Parallel components.

Caution should be taken in to account when using factor analysis for finding significant interactions amongst bi-polar characteristics. Since it may give misleading information and lead to wrong conclusions.

In other words, factor analysis is a determination of the smallest number of factors that can best represent the inter-relations amongst variables. In this thesis, the principle component method of extraction is used for the factor analysis. First Eigen values are derived which report the amount of variance explained by each factor. The determination of the number of significant factors takes place by Kaiser's Criterion and Scree plot (graphical presentation of the Eigen values). The Kaiser value should be 0.6 or above. And the correlation matrix should show at least some correlation of r=0.3 or greater. Tabachnick and Fidell's test may be used for the linearity test. Outliers can be changed to less extreme values or be removed. It is also important to investigate if a certain variable belongs to more than one factor. The method of rotation is used in order to maximise high item loadings and minimise low item loadings. There are two common rotation techniques: orthogonal rotation and oblique rotation. In this thesis the common rotation method of orthogonal Varimax is used. Orthogonal Varimax rotation first developed by Thompson [44] is the most common rotational technique used in factor analysis which produce factor structures that are uncorrelated [43][45, 46].

3.3.2 Regression models

Often, a scatter plot is used to illustrate the relationship between two variables. If a certain trend (linear or curve) is identified in the plot, it may suggest that the two variables correlate. Also correlation coefficients are used providing a numerical measurement of the linear relationship. The most well known statistic used for this purpose is Pearson correlation coefficient. This coefficient is ranged between -1 and 1with positive values indicating positive relationship and reverse. Application of this statistic is beneficiary, since the scale, location and units of the variables do not affect it. Linear models describe the relationship between two sets of variables; the dependent (the predictor) and the independent (the explanatory variable) [47]. Typical linear regression model with one variable is displayed below:

Equation 50

$$y = \beta_0 + \beta_1 x$$

Equation 51

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad i = 1,2, \dots, n$$

In which β_0 is known as the intercept and β_1 is the coefficient describing the relationship between Y and x whilst other factors are fixed (also named as the slope parameter).

[43] In contrast, oblique rotation produce correlated factors. In fact it produces more accurate results for research involving human behaviours. However, previous research such as Ou, have used the Varimax orthogonal rotation technique; therefore, in order to draw a controlled comparison between the methods, Varimax rotation technique is also applied here. Regardless of which rotation method is used, the main objectives are to provide easier interpretation of results, and produce a solution that is more parsimonious.

ε is the error term which indicates the effects of unobserved factors[44]. The errors may show some recognisable patterns which may suggest that other variables have significant affect on the response variable[45]. The linearity of the model indicates that a one-unit change in x has the same effect on Y regardless of the primary value of x. Method of least squares is used to derive an estimate of the variables. For regression models including more than one independent variables matrix approach to regression has been taken in MATLAB. Where Y is the vector of n observations of the response variable, β is the vector of coefficients, ε is the matrix of errors and x is the matrix of n variables (i=1,2...n).

Equation 52

$$Y_{n\times 1} = X_{n\times k}\times\beta_{k\times 1} + \varepsilon_{n\times 1}$$

Equation 53

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{12} & \cdots & x_{1k} \\ 1 & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

Multiple regression models are derived using iterative process[46] applied in three different ways; stepwise, forward and backward. Forward method is when one variable is added to the model one at a time. Each time model adequacy is tested by the F-ratio[47] test. Models with one independent variable are compared using the coefficient of determination ofR^2[48], and adjusted R^2 for more than variable. This process is repeated until the model consists of the variables that best describe the variation and adding more variables does not change the F-ratio or R^2 significantly. Additionally, another statistic is used which measures the model's mean square error relative to the full model's mean square error denoted as C_p[49]. The backward method is the same but starts with the model including all the possible combinations of the variables and takes a variable out on each step. The stepwise method

[44] The average value of u across the population is equal to zero.
[45] Since they are approximations of random errors and should not show any specific trend.
[46] By this, after the model formulation takes place, the data fitting of the model is criticised.
[47] F-ratio tests the null hypothesis of $H_0: \beta_1 = \beta_1 = \beta_2 \ldots = \beta_k = 0$. Also note that for single variable regression models, this test is equivalent to t-test with squared ratio.
[48] $R^2$is the proportion of the variability explained by the model (Multiple correlation coefficients indicated as R is the best linear combination of the explanatory and fitted values).
[49] Models with smaller C_p are preferred (such that $C_p \sim p$).

is a mixture of both but does not search all the possible combination therefore is not very advanced [47].

3.3.3 Participant agreement test

Participant agreement test is used when the nature of the experiments may cause participants to respond by chance, as the results are based on participant's subjective interpretation. Certain measures such as the Kendall coefficient of agreement u are used for similar purposes but only when the objects could be ranked and certainly not for this kind of data [48]. For determining the kappa statistic K; the coefficient of agreement for nominally scaled data, it is first required to produce a table for each bi-polar characteristic containing frequency values for all 28 colours. Taking heavy - light in Table 7 as an example; columns indicate the five different categories of -2, -1, 0, 1 and 2 and the rows contain the frequency responses for the heavy - light values for each of the 28 colours. The equation used for calculating the Kappa value is as follows:

Equation 54

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

where $P(A)$ is the proportion of times that the k raters agree and $P(E)$ is the proportion of times that k raters are expected to agree by chance. Complete agreement will result in K=1; and total disagreement will result in K=0. Like all other test statistics, the significance level of the K can be tested via the z value which is derived via $var(K)$ (considered as approximately normally distributed for large number of response).

Equation 55

$$z = \frac{K}{\sqrt{var(K)}}$$

Eventually, to test the hypothesis $H_0: k = 0$ against $H_1: k > 0$; it can be concluded that if the z value exceeds the ∝ significance level, then the agreement amongst the different colours can be considered as significant [48]. The algorithm for the Kappa test which has been programmed in MATLAB.

3.3.4 Kendall's coefficient of concordance

Kendall's coefficient of concordance *W,* is specifically used for when there are more than two sets of ranked data and reports the degree of association amongst them [48]. Kendall's *W* is measured by the formula below;

Equation 56

$$W = \frac{12\sum R_i^2 - 3k^2N(N+1)^2}{k^2N(N^2-1)}$$

where R_i stands for the sums of the ranks for each group, N is the total number of colours being ranked and k is the number of sets which are the bi-polar characteristics. X^2 is used to evaluate the significance of *W* via the chi-square distribution:

Equation 57

$$X^2 = k(N-1)W$$

The derived value of Kendall's *W* coefficient describes that the level of agreement amongst the rates.

4 Colour semiotics

4.1 Introduction

Colour is a powerful emotional signal in any graphical communication as it draws upon associations from the past and present [49]. The term signal, or signs, is of interest to the field of semiotics, also known as semiology. In this chapter, the term semiotics is studied and defined thoroughly, in order to open the doors towards a better understanding of one of the most important perceptual properties of colour. Thus, this chapter starts with a brief explanation of colour properties and design. Then, semiotics is referred to as one of the properties of colour and is discussed in the context of design. Understanding colour semiotics and design, leads to a need to understand the relationships between colour and affective parameters. Therefore, the research done by Ou and colleagues is described and critically analysed at the end of this chapter which is used later on as an inspiration for the deeper understanding of colour semiotic models.

4.2 Colour properties

Colour in the physical world, primarily results from the properties of pigments and dyestuffs that absorb and/or scatter light [14]. However, it is more than just a physical concept. The colour of an object can be shown to depend upon properties of the observer and properties of the background or surrounding colours [8]. As mentioned earlier, perceptual colour can be typically described by three terms: hue, colourfulness and value. In this section we review properties of colour that impact on human emotion, response and physiology. In particular, the properties are grouped under the headings, colour preference and colour harmony. Some researchers use the term colour emotion [50, 51] but it is considered to be a general term that encompasses some, but not necessarily all, of the three headings used in this review.

4.2.1 Colour Preference:

One of the earliest experiments to study people's preference for colour was conducted by Guildford and Smith [52] which revealed a preference for blue and green colours and a dislike of yellow; in addition, more saturated (colourful) colours were preferred and along with some gender differences[51] [53], bearing in mind that colour preference changes in accordance to context. The variables that influence the perception of colour on a product or packaging in the market place are not in full control of the designer or marketing

[51] Further on, the results of this study show agreement with this fact.

manager therefore, numerous colour preference studies have been carried out for particular products such as the food domain where the meaning of food is shaped much before the consumption. In particular, culture and psychological drivers play a significant role in shaping how we assign meaning even before tasting it. Studies have demonstrated that sensory experiences only account for 25-50% of variability [54]. For example, a study of consumer preference for the colour of chicken flesh revealed that consumers perceived the yellow colour of corn-fed chicken negatively [55]. Also, the colour of meat has been shown to be an extremely important factor that influences consumer purchase decisions (consumers reject brown meat) [56] and it has been reported that the average loss of meat sales due to colour deterioration was 3.7% in the USA [57]. Some colours are linked with specific tastes such as red and pink with sweetness, blue and white with purity and refinement, and green with mint [58]. From a practice point of view, the design process is often constrained by a combination of time, resource and expertise constraints in order to fully accommodate colour. This opinion is based on interviews with brand and design consultants and marketing managers in the course of personal and professional interactions. Thus, the usage of colour tends to be left to the aesthetic judgement of designers. Some large multi-nationals corporations, brand consultants, design agencies, and specialist research agencies are now paying attention to individual sensorial components of products (e.g.,[59]).

4.2.2 Colour Harmony:

The search for the rules of colour harmony has occupied the thoughts of some of the greatest artists and scientists [60]. There is no clear theory or definition of colour harmony. Judd described it in terms of two or more colours seen in neighbouring areas that produce a pleasing effect. Contemporary colour theories also referred to as colour theory in the field of art and design, have distinguished colour harmony as being monochromatic, complementary, analogous, and split-complementary[52][61]. There are of course many different types of colour wheel [62] and it is not clear which colour wheel would give rise to

[52] This has led to contemporary colour harmony theories (typically referred to as colour theory in art and design) that describes colour harmony as monochromatic (colours chosen with the same or similar hue on the colour wheel [Colour wheel is a circular presentation of the colour spectrum mainly used by artists and designers. Colour wheels can display from three or more colours.]), complementary (colours chosen with hues that are opposite each other on the colour wheel), analogous (colours chosen with quite similar hues on the colour wheel), split-complementary (three colours whose hues form a triangle in the colour wheel) etc.

the best predictions of colour harmony [63]. Moreover, it is clear that hue relationships alone cannot predict colour harmony [60]. It is, however, difficult to develop consistent and effective rules or models that can predict colour harmony since tastes change from generation to generation and according to an individual's age, gender, race, education and cultural background [64]. Indeed, some have argued that "it is quite evident that there are no universal laws of (colour) harmony" [65].

4.2.3 Psychophysical effect of colour properties:

There is evidence that colour can affect people's cognitive state and psychometric performance for various tasks. In a 1921 study Pressey [66] concluded that the level of colour saturation (controlled by lighting) had an impact on participants' finger tapping, multiplication and continuous reaction performances but related research [67, 68] did not support this. However, there is now a substantial body of published work that demonstrates the effect of colour on cognitive and psychometric performance in various tasks. For example, it has been shown that the concentration threshold for detection of sucrose is lower in a green aqueous solution than in a yellow solution [69]; blue- and green-coloured beverages have been shown to have greater thirst-quenching properties than similar yellow and red beverages [70]; the colour of political advertisements has a significant effect on subjects' judgement of the politician's personality characteristics [71]; the colour of ballot papers influences voters' choices [72]; the colours of walls affect the degree to which preschool children cooperate with each other [73]; the pink colour of walls can have an effect on the decrease of blood pressure which is used in prisons in Switzerland for fighting against aggression [74].

However, Wise and Wise [75] reviewed over 200 studies in environmental design to ascertain whether the use of colour in the design of interiors could be scientifically shown to have any effects on human performance, health or well-being. Wise and Wise did find some demonstrable effects that justified the use of certain colours in interior designs. However, they also reported that the notion that colour effects were 'hard-wired' and ubiquitous (for example, that certain colours would have a calming or energising effect), was too simple. In reality, colour can have strong effects but it can depend upon context, background (colour contrast being particularly important) and on the individual.

If colour affects the level of arousal felt by an individual [76, 77] then there may be a one-dimensional (inverted U-shaped) relationship between the single arousal dimension and affect [78]; however, others believe it is two-dimensional [79-86] relationship [87]. The question of the arousability of colour [88, 89] is of particular interest in the field of ergonomics and the design of office interiors. The results of some studies show, for example, that warmer colours (red, yellow and orange) have a greater arousing effect than other colours such as blue and green which are categorised as cool colours [88]; other studies [90, 91] have shown that saturated colours are more exciting[53][92-94].

Having colour properties defined, it should be included that the certain trends described about properties of colour vary depending on the colour source, for example; psychological and also physical effect of coloured light can be found in a different way compared to coloured dyes or coloured patches. Cases of research carried out on the effect of coloured light show that coloured light also has psychophysical effects which can be measured and even modelled [95, 96]. Another example of the research carried out on the colour lighting effect concludes that coloured light effect differently on people with different activity, evaluative factor and potency, also cyan, yellow and white were perceived as the most pleasant, comfortable and elegant coloured lights, blue and red on the other hand were perceived as agitating and heavy [97].

4.3 Design

Design in the commercial context is widely understood as the complex combination of creative and managerial process that overtly and covertly applies technical, functional, social, cultural, economic logic to objects[54]. “The etymology of design goes back to the Latin word de + signare and means making something, distinguishing it by a sign, giving it significance and designating its relation to other things, owners, users or goods. Based on this original meaning, one could say: design is making (sense of things).” [98]. Despite the rich, multi-dimensional and granular nature of “making sense”, it is commonly reduced to the binary activities of technological and functional innovation on one hand – normally considered the remit of empirical methods[55]; and styling innovation on the other – normally

---

[53] A study by Mehrabian and Russell considered the “importance of individual stimulus screening ability in the study of environmental factors and human performances”. They developed a questionnaire that identifies low, high and moderate screeners. Those who are distracted by external stimuli are called low screeners and the reverse is true for high screeners. The impact of colour in an office setting would affect individual’s differently depending upon whether they are low or high screeners.

[54] For the purposes of this research, objects will be referred to as artefacts.

[55] For example industrial engineering and design.

considered the remit of aesthetic subjectivity[56]. In reality, however, the distinctions are blurred as various stakeholders (technologists, engineers, designers, managers, consumers) influence and shape technology, usability and aesthetics. An entire field of sociology is dedicated to the social construction of artefacts [99].

Contemporary international efforts of industry bodies, governments and academia revolve around the idea of "good design". These efforts are aimed at finding methods and approaches to technological and creative development with the final objective of design innovation. The tension between purely functional developments of design versus creative expression is not a recent improvement. Product designers trained in Europe often espouse Dieter Rams' design principles: "integrity in design accompanied by true functionalism rather than style" [100]. He defines "good design" as innovation, aesthetics, unobtrusive[57], honest, long lasting, thorough and environmentally friendly whilst making a useful, understandable product by using as little as possible. However, design is not a straightforward process involving only science and creativity. Designs are created in an industrial setting, which normally follow a series of steps. The commercial imperative demands that a consumer need be converted into a viable product or design solution with initial stages of needs identification and definition, research, brief, concept, development and presentation [58]; brief explanation of each stage is displayed in Figure 14 .

[56] For example graphic design, fine arts.

[57] Products that are considered to be tools are neither decorative objects nor works of art. So an unobtrusive design for such objects would be a neutral design leaving more room for the user's self-expression.

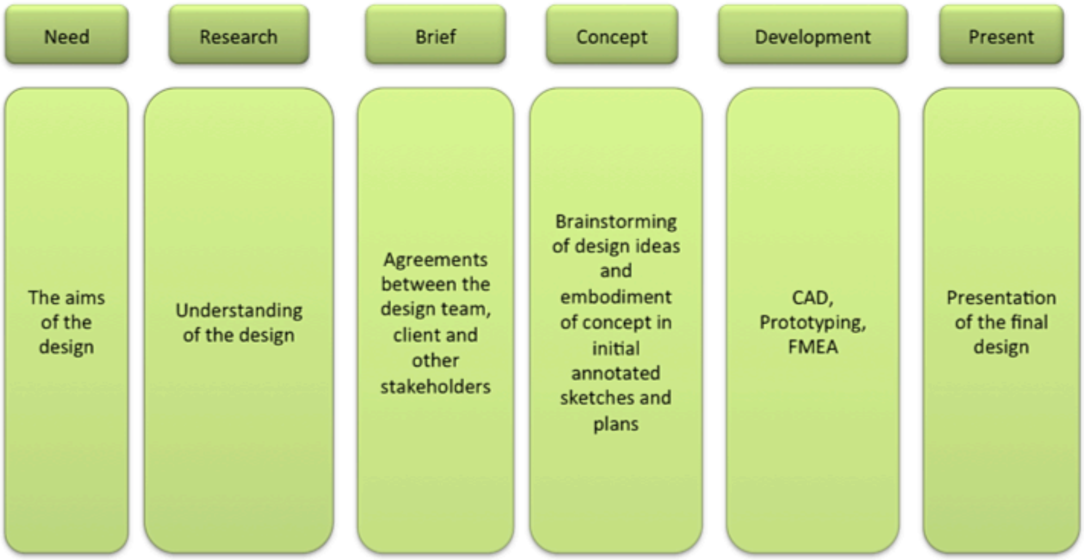


Figure 14: A design process.

The design process involves a very high level of creativity which functions parallel to the design problems and aims. Initially product is defined as the combination of physical characteristics and service elements that are set according to the customer's need [58]. It is clear that the consideration of audiences, industrial sectors, product categories, competition and brand positioning leads design towards specific strategies and interventions in successful design. Therefore, the design process is highly dominated by the target market satisfaction factor. However, creating an appealing design requires understanding of branding, colour and language [58].

However, there is a lack of theoretical and methodological consensus on what constitutes appropriate levels of understanding of consumer needs. Ulrich [101] have prescribed a methodology for product concept development, which recommends identifying both explicit and latent or hidden needs . However, the prescription to understand the user is directed at sensitising the *"use environment*" of the product. Though Ulrich *et.al.* suggest use of qualitative methods such as interviews, focus groups and observations, their formulation is underpinned by a structured approach that entails users articulating or demonstrating their problems with a product or expressing their needs and wants. The opportunity to uncover opportunities in latent needs is left to any "surprising" mentions by users. Despite their empirical insistence, Ulrich et al end up prescribing interpreting users' articulation, and illustrate examples of right versus wrong interpretations. Indeed, Shiba *et.al* point out that making clear the vague customer requirements is acknowledged to be one of the weakest parts of the product development process [102].

Baxter prescribes culture as one of the highest levels of perceptual determinants of product style in a mix containing many different levels of determinants. The process of getting the mix right is a complicated one and Baxter argues that many manufacturers approach design consultancies after technological development leading to superficial design exercises in styling through colour, pattern, form etc., rather than making design thinking the very basis for product development. Baxter recommends four faces of attractiveness, one of which again points to culture, that is the symbolic meanings generated by product design. Baxter argues that evidence in product design demonstrates a direct link between the design's socio-political background and the resultant objects. For example, the communist Soviet-Union frowned on the ruling classes' lavish and ostentatiousness styles and this led to the emergence of strong utilitarian and industrial themes. The post-communist Russia has witnessed a boom in the luxury trade market, with uncanny resemblances of the present concentration of wealth and power in the hands of a few with the past.

Baxter's book, praised for its integrative approach to product design, focuses on principles of product styling and creativity such as gestalt rules, golden ratios and sections, bio-sociative attraction, social/cultural/business effects, attractiveness amongst others but does not delve in the details of elements of design such as colour [103].

Colour is a very important component of brand-building, for e.g., logos, packages, and displays [104], it is known to be the central tool in the agenda-setting function of advertising [105]. Milton Glaser (1985) [106], the celebrated graphic designer, argues that despite no difference in costs to quality packaging, generic packaging – which often uses very little colour – is intended to produce the impression that "no time was spent doing it and no cost". Thus, generic design, which uses minimum – or restrained – colour is a calculated response to a culturally aware consumer who realises she/he does not need the fancier product. Thus, Glaser acknowledges that packaging designs and their constituent elements constantly reference explicit and implicit meanings in culture.

Colour has an attention-drawing property which operates as a visual noise, drawing attention towards sensory data, whether relevant or [107-109]. Colour holds vital meaning and if we limit ourselves to a two-tone communication (black and white) it would be poorer in information and meaning [110] (where specially in adverts the communication would be partial). At worst, it could lead to misleading communication. Consequently, adverts that employ colour for its semiotic functionality impacts on humans at a level further than denotation as it creates association between the product and some other good (a pleasurable, generally positive, occasionally negative, but always purposeful) message [105].

4.4 Definition of Semiotics

Colour is a powerful emotional signal in any graphical communication as it draws upon associations from the past and present [49]. The term signal, or signs, is of interest to the field of semiotics, also known as semiology. The word Semiotics can be traced back to the Greek word sēmeîon[58][111]. Contemporary usage of the term is associated with two scholars, the Swiss linguist Ferdinand de Saussure and the American philosopher Charles Sanders Pierce, who are considered the co-founders of the discipline. Saussure described semiology as a science that studies the relationship between signs and social life, which in turn influences social psychology or the circulation of social meanings. Pierce, on the other hand, used this term to elaborate on a 'formal doctrine of signs'. He argues that human beings 'think only in signs'. For Pierce, signs can take any form – words, images, sounds, odours, flavours, human gestures or actions, objects. These signs are empty until human beings invest them with meanings [112].

The importance of semiotics to the study of colour is underscored by its ability to mediate the link between the external physical world and the internal construction of the same, which is shaped by culture. Kenney argues that colour as a sign could be given out or created by humans, plants, animals or inorganic matter. Thus, for Kenney (2005) colour is a sign that mediates humans' internal experience of the external physical and symbolic worlds [113]. Though Kenney intellectual interest is in signs and representation, we could apply his logic of looking at a photograph and ask the question: When YOU look at a paper print of the colour red, what do you see? Do you merely see a paper print of the colour red? Are you reminded of something that is red in colour, for example a favourite dress or car? Does it trigger the memory of an experience? Do you relive the experience? Are you filled by a certain emotion? For Hyman (2000), colour serves as one of the elements of a

[58] 'Sign is described as a meaningful unit which is interpreted as 'standing for' something other than itself'.

configuration that help humans to understand the representation of a physical object or a cultural idea [114].

Signs work in subtle and sophisticated ways within a culture. For example, in the UK, use of paper cups or serviettes in a Michelin-starred restaurant would be considered “cheap”, while it would be perfectly acceptable in a high-street coffee shop. This tells us about the usage of paper as a material within the culture of dining in a country. We can postulate that every culture weaves a web of meanings that are expressed through signs and symbols, which may or may not be fixed. Thus, signs mean little when stripped of human and the cultural context in which they are interpreted. This interpretation should also consider meanings in the context of their histories as well as the flux at the time of its interpretation. This allows correct interpretation of a sign. For Eco (1996), human perception and understanding of colour is mediated by language in which the visual identification interacts with discrimination of colour allowing a chain of interconnected and interchangeable references within a given culture [115].

Colour classification, like all other forms of classification, is a mechanism for humans to cope with the physical world they encounter. Thus, though colour distinctions are naturally based, the natural distinctions, which are expressed in words, are culturally formed [115]. When the ability to discriminate produces subtle classifications, humans acquire a more powerful world view, changing the (semiotic) codes. Eco illustrates the role of language in limiting or expanding human understanding physical phenomenon by citing the example of how the mouse and rat are distinguished. In Latin there exists only one term, “mus”, while the English classifications are mouse and rat respectively. In Italian, despite the existence of two terms of ‘topo’ and ‘ratto’, the common usage is limited to ‘topo’ thereby limiting the average Italian’s ability to discriminate between the two producing a number of sanitary and social consequences.

Eco argues that the generally low colour discrimination levels produced by Farnsworth-Munsell test are an indicator of the limitations of language and culture. For example, Hindus consider red and orange to be one pertinent unit, the Latin did not clearly distinguish between blue and green, while Russians categorise blue into different portions, *goluboj* and *sinij*. Through a demonstration of the organisation of colours amongst Hanuńoo people of Philippines, Eco argues that the pertinentisation of the colour spectrum depends on symbolic principles, which may be produced because of practical purposes. Thus, for the Hanuńoo, four fundamental colours express the cultural principles of *Malagati* (light, weak, dead, pale), *MabiRu* (dark, rotten), *Malatuy* (wet/fresh/succulent) and *Marara* (dry, dessicated). The community finds a link between the four types of colours and nature, that is lightness and dryness are viewed as a linked unit while darkness and wetness are viewed as another.

Semiotics as a field of study is interested in signs, signifiers and the signified. The signifier is the actual material of the sign (for example the physical material of blood that denotes the sign "red" and "blood") while the signified refers to the conceptual meaning or meanings generated by a sign and its signifier. For example, blood as a signifier refers to the sign "red" and "blood", which in turn could generate a host of signified including injury, murder, violence (including its time), animal, human, danger, threat. Due to the nature of the signifier (blood) the sign "red" will probably not be invoked and in turn the other signified associated with red, for example passion, energy etc., would probably not be invoked. One of the broadest clarifications for semiotics has been made by Umberto Eco, who states: 'Semiotics is concerned with everything that can be taken as a sign' [116]. Thus, it helps us to understand the indication of signs and their meanings. It should be noted that signs are classified in to three main categories; icon, symbol and index. Iconic signs relate to their object in some kind of a visual resemblance (e.g., a photograph or painting of a leaf or a plane). Symbolic signs generate meanings through learned conventions (e.g., the word leaf), while indexical signs refer to their origins through some kind of directly physical or material connection (e.g., fossil of a leaf is an indexical sign of a leaf [112] (see Figure 15).

Figure 15: From left; symbolic, iconic and indexical examples for leaf and airplane.

Semiotics is a relevant framework for the investigation of colour and for the purposes of generating useful explanations of their usage and understanding. Semiotically, colour as a sign and signifier can generate different meanings depending on what kind of sign it is – whether it is indexical, iconic or symbolic.

4.5 Colour semiotics

A body of literature has emerged around the semiotics of colour: Firstly, colours as words serve as symbols whereby they represent cultural ideas. This link between colours as words and their cultural meanings is referred to in literature as colour naming. There is a significant body of literature on colour naming from sociological [67], psychological [68] and technological [66] perspectives. However, this paper is more interested in what colours signify, represent or communicate. The generalities of colour semiotics have been described by Caivano [117] who notes that there are three levels of representation; the sign that is doing the representing, the object/concept that is being represented and the observer who is interpreting the sign. It is clear that these levels apply to colours. Contemporary semioticians use a similar interpretative framework, describing the process encoding and decoding. Decoding involves the interpretation and evaluation of the meaning with reference to the relevant codes. Codes are systems of related conventions for correlating signifiers and signifiers in certain domains. However for visual perception; there is some disagreements about perceptual codes[59][7]. We refer to colours in our everyday communication to express our feelings. For example, we speak of “being green with envy” or “being so angry that we see red”. In addition, we characterise people’s behaviour and mental status using colours; for example yellow stands for cowardice, blue for depression etc. [105]. Many researchers have tried to affix meanings to colour [118] and certainly in many cases there is close agreement about the meanings of colours. For instance extrovert, danger, energy, war, strength, power, determination as well as passion, desire, and love are associated with red, while green is associated with growth, harmony, freshness, safety and fertility and blue is considered to indicate depth and stability while symbolising trust, loyalty, wisdom, confidence, intelligence, faith, truth, and heaven [119-123]. However, it is important to understand that such agreement may only be valid within a specific society or

[59] The concept of perceptual codes suggests *non-isomorphism* (refers to the de-similarities of two different structures) between mental representations and sensory data (*mind-world*). Some semioticians regard sensory perception as a code. Note that in the notion of perceptual code intentional communication does not exist in other words there is no need for a sender.

culture and may change over time. For example, while red is associated with danger in the Western convention [124, 125] in China red is associated with good luck [126]. The cultural study of colour is a complex and fascinating topic [70].

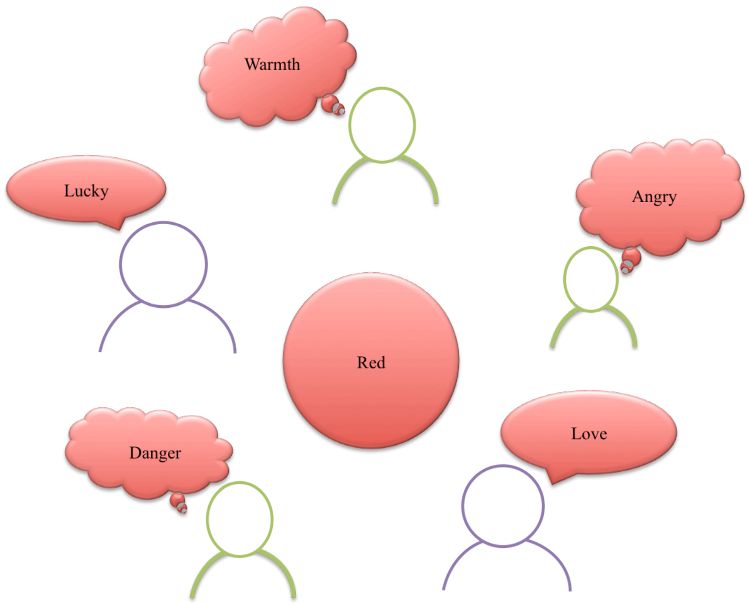


Figure 16: An example of different individuals' diverse semiotics

There may be evidence that some colour semiotics have more global agreement; for example, the notion that certain colours are warm and others are cool [88]. It is likely that there are at least three sources for colour semiotics, these being innate (e.g. we may be born with a predisposition to make certain colour associations such as the link between red and danger), socio-economic (e.g. the representation of purple with royalty and richness in some cultures is certainly linked with the fact that purple dyestuffs were once more expensive than gold gram by gram) and socio-cultural (e.g. the use of red envelopes to give money to people in China has probably resulted in Chinese people associating red with good luck). According to Hupka's categorisation [117] of emotions, the first of these (innate colour meanings) would be classed as primary emotions whereas the other two would be classed as compound emotions [127-129] (see Figure 17).

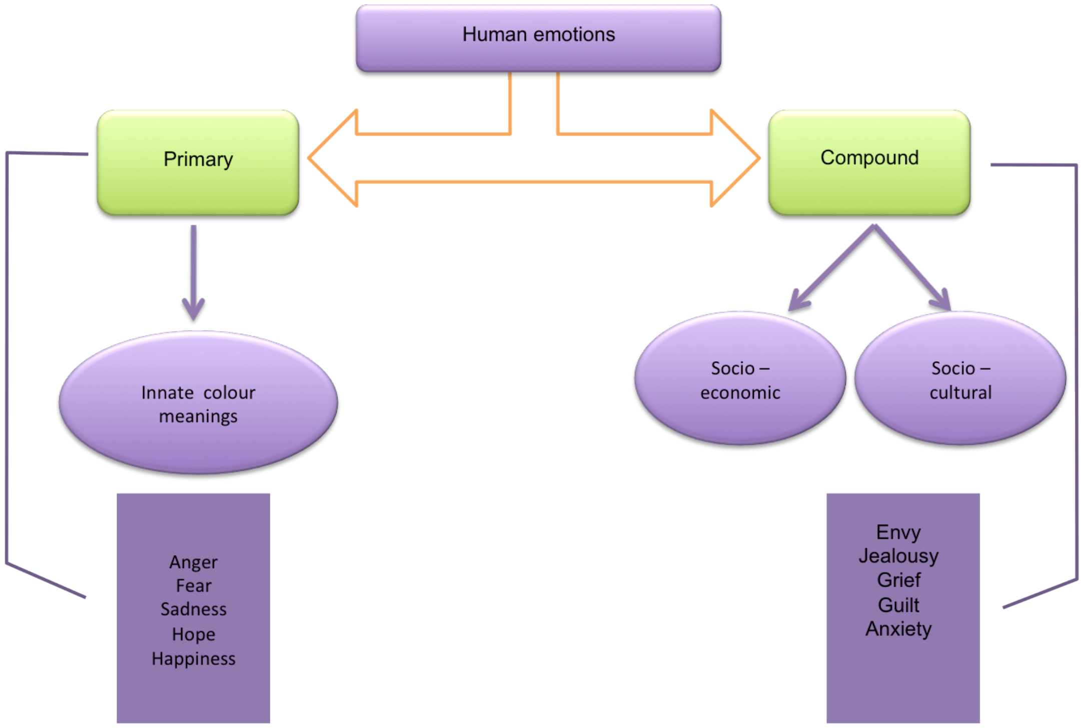


Figure 17: Categorisation of colour semiotics using Hupka's framework of primary and compound emotions.

4.5.1 Colour as an indexical sign

Indexical signs bear a direct link to what the sign refers to (the referent). For example smoke is an index to fire; a sneeze is an index of a cold. In the context of colour, for example, an orange coloured drink would normally serve as an index to orange squash and usage of any other colour would generate disbelief in the minds of consumers. Recently, however, celebrity chefs and some restaurants have attempted to challenge this link between the visual and taste sensorial of food by serving food in a dark environment or even adding very odd colours such as blue to potatoes. The FMCG industry however, continues to rely on the indexical link between food and its colour as these experiments have not gained popular acceptance.

4.5.2 Colour as an iconic sign

Colours can generate resemblance to objects and thus serve as iconic referents to them, and the link lies in the colour's connotation of the object. The colour red, for example, is an iconic representation of blood. It is important that the colour's attributes (e.g. hue or saturation) resemble the qualities of the object. In terms of design, the application of this colour requires additional signs in order to generate the iconic meaning of blood. For example, the illustration of the colour as a viscous liquid against skin would normally generate reference to blood for the audience. The understanding and use of colour, particularly in product or packaging design, is complex as it is one among the many signs that is deployed to persuade consumers about a brand's meaning.

Figure 18: The colour red used in an advert for Mini refers to blood by working with other signs – True Blood TV series, the subject of which is vampires and the use of the word "fangs".

The colour red links to blood through the denotation of the words "True:Blood" and the latter's link with the TV series True Blood, which is a TV series about vampires, and properties of the product – a liquid. However, the generated meaning is not literally drinking blood (though some consumers may not like that association) but through the brand value that is created by the associations of the brand with symbolic meanings that the interaction of the signs generates. One set of meanings generated is the thrill of nocturnal "prowling" and "preying" and consuming a forbidden object. Thus the usage of the colour red for the desired effect requires first an iconic fit (through hue and saturation) and the interaction with other signs.

Figure 19: An example of the colour red linked to blood.

4.5.3 Colour as a symbolic sign

The third type of sign is the symbol – a "word, hypothesis, or argument, which depends on a conventional or a habitual rule". A symbol is considered pragmatic, as both the sender and receiver of the text are aware of the meaning and it helps interpretation and communication. Symbolic signs serve to help action or thought. Furthermore, Pierce argues that there are three types of interpretants – the immediate, the dynamic and final interpretant. Using the previous example of the colour red, it requires a precise collection and organisation of related signs for it to be immediately interpreted as blood [112].

It is also important to note that cultural differences and type of product can have a direct influence on the way the colours are symbolised, taking the red for example in the food domain; In the West red foods are considered as highly desirable since they indicate passion and prohibition (Adam's apple). Whilst in far East (specifically China), red is associated with various organs of the body, therefore; red food such as sweet potato, carrot, tomato, red chilly are believed to have good impact on heart functioning and blood flow by removing heat from blood. On the other hand, in India red is associated with foods that are related to virility, vigour and festivity. Considering this fact does not only limit to the food domain, but rather all products and also various shades of the same colour do not symbolise the same thing and are also dependant on cultural contexts.

In general the nature of the product that the colour is applied to is crucial as well. For example in

Figure 20, red is used prominently within healthy yoghurts. A powerful tension is created between red (suggesting intensity, energy, indulgence), green and white which suggests wholesomeness, purification and restraint. This codes healthy eating products as miraculously offering indulgence without the penalties associated with it. Whilst in other products such as sweets (see Figure 21) red suggests childish base needs, energy and frivolity. It is also used to accentuate graphic elements that similarly suggest disorder.

Figure 20: examples of the use of red in yoghurt packaging.

Figure 21: Application of the colour red in sweet packages.

The different symbolisations that exist between the cultures can root back to the definitions and notions which have been established during the history. In western antiquity for example, Aristotle postulated that the colour red, green and purple were the only unmixed intermediate colours. For the Romans on the other hand, red was the colour of light [130, 131][60]. Red in Indian-Hindu cultural traditions is the colour of matrimonial union, marriage, divinity, fertility, heat, erotica, passion and energy. The matrimonial symbolism of the colour red was to a large part the preserve of Northern Indian cultural traditions. On the other hand, historically the colours of matrimonial union in South India are largely white (borrowing from codes of purity). However, with the explosion of mass media in India, the north Indian culture is permeating all parts of Indian society and the colour red is beginning to be codified as the colours of matrimonial union across India. However, red in Chinese culture is associated to femininity, good luck, happiness, marriage and success.

4.5.4 The advantage of acknowledging colour semiotics

Although explained earlier, it is important to highlight that a product which is well designed can grant a firm with a competitive benefit that will enable it to achieve a higher market share [132]. Immediately after observing a product we derive a special meaning from it, which is developed from the complement of two stages;

1) When the appearance of a product is observed by a consumer, specific physical elements such as colour, shape and texture are perceived which together make the general concept of the design of the product.

2) Specific combinations of colours, materials and other physical aspects give the product a specific look that can be described by a certain characteristic of its appearance [133]; for example, modern, metallic, frilly, simple, sleek, old fashioned etc.

Those characteristics mentioned in the second stage are more theoretical than the separate physical features mentioned in stage one [134-137]; this shapes the importance of the derived characteristics from the appearance of a product (colour semiotics). Colour is also known as a marketing cue which global managers use [138] to sustain, create and adjust brand images in customer's minds [104]. The meanings which are associated with different colours are the communication tools for brand image and that's why it is very important to marketers [139]. Thus, if a company is not able to make the right connection between a

[60] Gage argues that the colour red was associated with divinity as evidenced in its usage in Greek temples (temple of Aphia in Aegina) as well as temples in India (Taxila).

certain meaning and the appearance of a product through a reasonable design, then its chance of success amongst the market is decreased. Initially, it will not be able to form a competitive benefit [140-144]. In other words, the wrong use of colour semiotics results in the wrong response from the consumer which will decrease the market's interest. The reason is that consumer choice takes place mostly around a product which is more pleasurable to use. This feeling only arises from the appearance of a product which motivates by its functionally, artistically, symbolically or ergonomically design [145]. The process of designing a product and to prepare it for mass production is a lengthy and time-consuming procedure. Markets frequently change their taste for colours and that is why it is important to be able to predict a market's preferred colour for a given product (colour forecasting) [146]. However, a designer can never be sure that the derived meaning of his/her design can be the same certain objectified meaning as he/she wants to be derived. This idea has been postulated by Krippendorf (1989)[98] who noticed that because consumers don't automatically get the intended meanings out of the product appearance, companies are pushed into exchanging the meanings of the product through marketing campaigns which are not cost effective. On the other hand, the gap between designers and consumers gets wider when pointing out there opinions about the attributes of a product. Consumers judge product attributes based on their own experience and what they feel about that specific design. They have less knowledge and see less differences between objects than professional designers [147, 148] which according to literature this difference between specialists and non-professionals can be understood by observing and evaluating a great variety of stimuli [147-149]. On the other hand "marketing a product is *re-active* and not *pro-active"*; consumer interaction is based on existing colour palettes and not very much on the unknown colours in the upcoming future [150]. Fortunately, according to Blijleven's (2009) [145] research results, a product's characteristics are universal amongst the consumers and the products and to be exact modernity, simplicity and playfulness present closer attributes to consumer perception than what is portrayed in the literature produced by artificial stimuli [137, 151-153].

The importance of colour is evident even when it has only been changed on the same product (e.g. change of the colour of a fizzy drink can from green to orange) resulting in a sudden rise or fall in its sales. James Mandle, who worked as a colour consultant, changed the light blue colour of Ty-D-Bol's cleanser bottle to stark white letters on a dark background so that the new bolder colours promote strength and cleanliness, which gave a sudden rise of 40% in the sales in only 18 months. Bearing in mind that it is not always required to change the hue as changing the chroma or value will create a new experience while still maintaining the actual hue associated to that product [87] (e.g. Ritz Cracker advertising and packaging 1980[61][154]). Some producers do this as it is significantly cheaper to change the colour of a product than designing it all over again [150], but others do this as a way of creating new feelings for the consumer. Taking the fizzy drink can as an example, most of the people who experienced tasting both cans thought that the drink inside the orange can was much sweeter than the drink in the green can, while they were made from the same recipe. So this is why derived attributes of a product vary by different colours of a product because they form alternative emotions. Bearing in mind that sometimes the novelty of a colour change is the reason for its rear impact (e.g. pink prison experiment from the late 1970s which concluded that the novelty of the colour change had generated the less aggressive effect for the prisoners locked up in a pink room [155]). Time For this reason we see that many brands keep their colours specifically to themselves for the sake of their branding strategies like Cadbury's purple or Kodak's yellow or BP's green and use colour as their powerful tool [156] which can cause good and positive feelings, as well as favourable disposition towards their brand [157] . In conclusion we understand that when a colour is used for a product, it can not cover an unfinished or ill-conceived strategy, but when it is used as part of a good design or strategy, it can contribute to a more successful marketing effort [158, 159].

4.6 Colour semiotic models

Over the past few years many colour scientists have been investigating the possible relationships which exist between colours and essential factors such as age, gender, culture. Up to now, not many of the researchers have been very successful in modelling these relationships, because of the experimental restrictions which limits the diversity of respondents. Also, the relationship between colour semiotics and colour variables such as the LCH values has become a subject to research such as the so called colour-science-based colour emotion models by Ou [1], Sato [160] and Xin and Cheng [161]. These different models show significant agreement together, however; they suggest that the colour emotions (colour semiotics) are independent of culture and nationality. The most commonly used model amongst the colour scientists is Ou's colour emotion which is why it

[61] In early 1980 Ritz Cracker advertising and packaging used a higher chroma degree while keeping the same hue, making it effectively "ritzier".

is the main focus of this study and is used and compared throughout this research. Across all the existing colour models, bi-polar characteristics maintain the main foundation as their relationship to colour values are of interest.

4.6.1 Bi-polar characteristics

In colour research bi-polar characteristics (colour emotion scales) are used so that colours can be emotionally or logically attached (assessed) to one of the word pairs. In earlier research many bi-polar characteristics have been studied which have made the research process lengthy and sometimes not possible, which this is why earlier research on colour has been concerned about categorising bi-polar characteristics in to groups and also reducing them in to smaller numbers (by using the factor analysis[62])[43, 162, 163] (see). For example, 48 bi-polar characteristics were grouped in to six main categories of happiness, showiness, forcefulness, warmth, elegance and calmness by Wright and Rainwater [164]. They also found a certain relationship with the colour attributes as the influence of lightness and chroma was greater than hue. In another study; 12 bi-polar characteristics were categorised in to four factors of impact, usualness, evaluation and warmth by Hogg [165] whilst in a study for simulated interior spaces, he found five factors of dynamism, spatial quality, emotional tone, complexity and evaluation [166]. In both studies complex relationship with colour attributes were found except for warmth, dynamism and emotional tone (from the simulated interior study) were they had clear relationship with chroma and hue. Later on Kobayashi [167] and Sato [160] found similar dimensions of warm-cool, soft-hard (potency) and clear-greyish (activity) which were found to be associated with hue, lightness and chroma. And finally, Ou has used 10 bi-polar characteristics that are mostly applied by previous studies.

It is evident that similar factors have appeared across these studies which can show that these factors are global no matter what the main bi-polar characteristics are. However, some argue that the appearance of similar factors across different studies is because of the similar bi-polar characteristics being used [168]. This can be rejected to some extent, regarding to the nature of factor analysis. This takes all the bi-polar characteristics in to account and also despite the different bi-polar characteristics, the initial factors are the same marking a universal status for them [169].

---

[62] Explained in section 3.3.1 Factor analysis.

4.6.2 Ou's colour emotion model

Ou has series of work based on the colour emotions and colour preferences models based on single and multi colours. For the purpose of this research, part one of his study, which is focused on single colours, is investigated. 31 observers were involved in his psychophysical experiment where they are classed as 14 British and 17 Chinese and are proposed with 20 coloured patches chosen from the NCS atlas and distributed in the L*a*b* space as shown in Figure 22. The 3x3 inch colour patches were presented randomly one by one to each participant, illuminated under D65 light in a veri vide viewing cabinet (with a uniform grey background with $L^* = 50$) and displayed in a distance of 40 inches from the observers eyes. The colour values were then measured by the Minolta CS-1000 tele-spectroradiometer, located at the observer's eye position. Each participant had to assess each of the colour patches to the 10 bi-polar characteristics (as described in Table 2) orally and in one of the British or Chinese language (depending on relevant nationality) which was defined individually by the Cambridge Advanced Learner's Dictionary before the experiment. The data collection method used was the paired comparison method; this means that each participant had to choose one word between the bi-polar pairs for each colour, for example; the participant was presented with a colour say red, and had to respond orally if it was heavy or light, soft or hard, warm or cool and so on. The frequency values derived from the data collection were converted in to z scores which are based on the Case V of Thurstone's law of Ccomparative-Judgment [42]. Large positive z scores indicate higher percentage of observers choosing the first word whilst negative z scores show that mostly the second word was chosen and of course zero indicates that half of the participants chose the first word and half chose the second. More information about this methodology has been described in scaling methods section of this chapter.

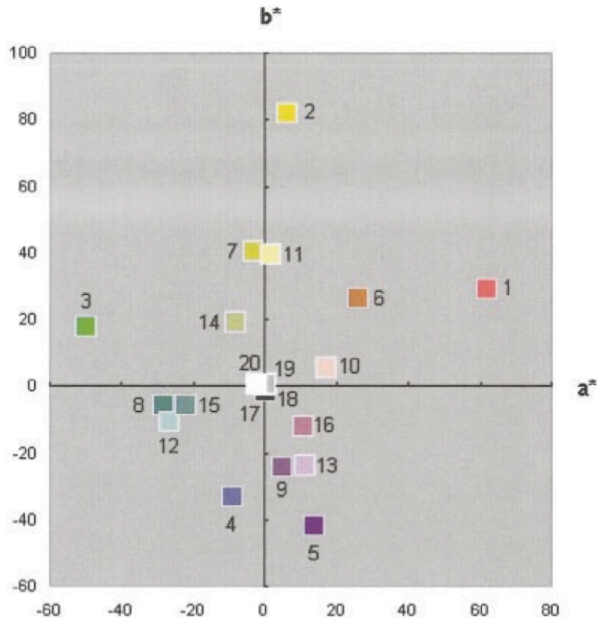


*Figure 22:Ou's colour samples in the L*a*b* space [1].*

Table 2: Bi-polar characteristics used in Ou's study.

Primary factors	**Bi-polar characteristics**
Evaluative factors	Clean – dirty
	Fresh – stale
	Like – dislike
Potency factors	Heavy – light
	Hard – soft
	Masculine – feminine
Activity factors	Warm – cool
	Modern – classical
	Active – passive
	Tense - relaxed

Ou used different methods in order to find the relationship between colour attributes and bi-polar characteristics. He used the observer accuracy value; which indicated the extent to which accuracy was poor or not (according to the wrong decisions (w_i)).

Equation 58

$$Accuracy = \frac{\sum_i (1 - w_i)}{N}$$

In order to find the effectiveness of gender and culture parameter he used the Pearson product moment correlation coefficient which results signified little gender differences and also cultural differences amongst a few bi-polar characteristics such as tense-relaxed and like-dislike. Also, factor analysis was applied for the investigation of cultural differences; which was applied on two groups of British and Chinese separately. Both resulted in three separate factors which described 79 percent of the variance for British observers and 90 percent for Chinese. Interestingly, British and Chinese factors had the masculine-feminine, hard-soft and heavy-light in one factor, modern-classical, fresh-stale and clean-dirty in another factor and warm-cool in the third factor. This similarity of factorisation between the groups suggests that they act in a similar way. Although, on the other hand, tense-relaxed, like-dislike and active-passive fall into different factors which suggest some level of differences. Three-dimensional plots of the factors were used for visualisation of this fact. Another approach was the development of the three-dimensional colour emotion space for single colours. By this, colours are placed in a uniform space with axes indicating the three factors extracted from the factor analysis (British and Chinese). This three dimensional space has been formed using component analysis on and coordinate determination and was formed on the basis of factors labelled as colour activity, colour weight and colour heat. An interesting finding in this research is the similarity of factors with Kobayashi [167] and Sato's [160] colour image scales. It is pointed out that although these studies were carried out in different parts of the world (Japan and UK), the underlying dimensions of colour scales are the same.

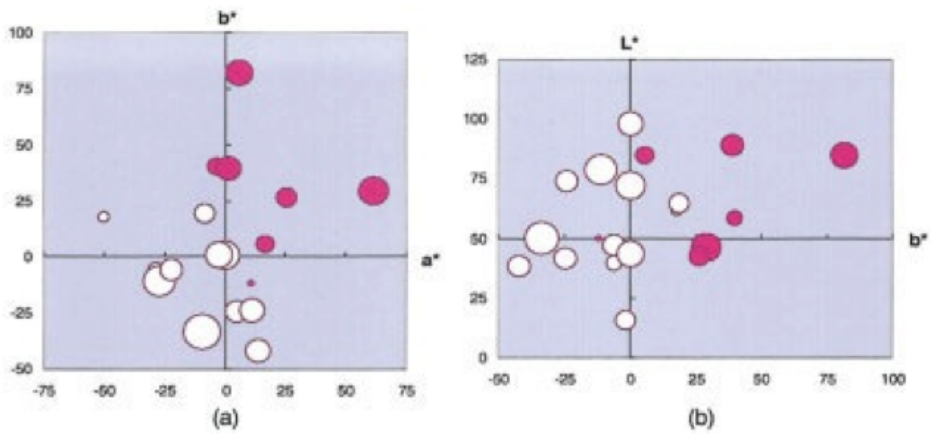


Figure 23: Ou's warm–cool in: (a) a-b* diagram and (b) b*-L* diagram. Red bubbles represent warm colours and white bubbles represent cool colours [1].*

The quantification of the colour semiotics (Ou refers to as colour emotions) is done through a bubble plot (Figure 23). Each bubble represents a colour, its location on the plot is based on the attributes of that colour, the colour of the bubble represents the chosen characteristic (e.g. red for warm and white for cool) and the size of the bubbles are associated with the intensity of that characteristic (e.g. big red bubble indicates that the colour has been mostly associated with warmth). The location and distribution pattern of these bubbles describes the relationship between the bi-polar characteristic and hue, lightness or chroma. The formulisation of this relationship has been done using the cosine of the angle between the colour and a test colour as followed:

Equation 59

$$WC = -0.5 + 0.02(C^*)^{1.07}\cos(H - 50)$$

Equation 60

$$HL = -2.1 + 0.05(100 - L^*)$$

Equation 61

$$AP = -1.1 + 0.03[(\Delta C^*{}_{N5})^2 + (\frac{\Delta L^*_{N5}}{1.5})^2]^{1/2}$$

Equation 62

$$HS = 11.1 + 0.03(100 - L^*) - 11.4(C^*)^{0.02}$$

Where L^*, H and $C^*$are the CIELAB lightness, hue and chroma and also $\Delta C^*{}_{N5}$ and ΔL^*_{N5} are the chroma and lightness differences with a test grey of $L^* = 50$. He also found the following predictive models for the three sets of factors;

Equation 63

$$Colour\ activity = -2.1 + 0.06[(L^* - 50)^2 + (a^* - 3)^2 + (\frac{b^* - 17}{1.4})^2]^{1/2}$$

Equation 64

$$Colour\ weight = \ -1.8 + 0.04(100 - L^*) + 0.45\cos(H - 100)$$

Equation 65

$$Colour\ heat = \ -0.5 + 0.02(C^*)^{1.07}\cos(H - 50)$$

Where a^* and $b^*$are the CIELAB coordinates of the colour. These models are compared with Sato and Xin's models in the next section.

4.6.3 Comparison of the existing colour semiotic formulae

Table 3 contains the formulae developed by Sato [160], Xin and Cheng [161] and Ou with the pereformance value of R^2 (which is derived by the calculation of the same twenty colour values used in Ou's experiment).

Table 3: Comparison between Ou's models and the performance of Sato [160] and Xin and Cheng's [161] models.

	Model	R^2
Sato	$WC = 3.5[\cos(H-50)+1]B-80$	0.69
	$HL = -3.5L^* + 190$	0.76
	$AP = ([0.6(L^*-50)]^2 + \left[4.6\left(1-{}^{\Delta H_{290}}/_{360}\right)C^*\right]^2)^{1/2} - 115$	0.59
	$SH = ((3.2L^*)^2 + [2.4(1-\Delta H_{290}/360)C^*]^2)^{1/2} - 180$	0.53
Xin – Cheng	$WC_{0°\ll h\ll 180°} = 0.154L^* + 39.378C^{*(0.372)} - 0.303H - 113.855$	0.45
	$WC_{180°\ll h\ll 360°} = 0.355L^* + 23.476C^{*(0.429)} - 0.159(360° - H) - 105.710$	
	$HL_{0°\ll h\ll 180°} = -3.340L^* - 0.476C^* + 0.037h + 175.467$	0.81
	$HL_{180°\ll h\ll 360°} = -3.477L^* - 0.264C^{*(0.429)} + 0.072(360° - H) + 182.866$	
	$DyPa_{0°\ll h\ll 180°} = -0.296L^* + 3.162C^{*(0.931)} - 0.07H - 68.835$	0.52
	$DyPa_{180°\ll h\ll 360°} = -0.120L^* + 4.385C^{*(0.864)} + 0.032(360° - H) - 84.791$	
	$SH_{0°\ll h\ll 180°} = 2.9L^* - 0.51C^* - 0.051H - 146.7$	0.5
	$SH_{180°\ll h\ll 360°} = 2.953L^* + 0.424C^{*(0.429)} - 0.02(360° - H) - 159.795$	
Ou	$WC = -0.5 + 0.02(C^*)^{1.07}\cos(H - 50°)$	0.74
	$HL = -2.1 + 0.05(100 - L^*)$	0.76
	$AP = -1.1 + 0.03[(\Delta C^*_{N5})^2 + (\frac{\Delta L^*_{N5}}{1.5})^2]^{1/2}$	0.75
	$HS = 11.1 + 0.03(100 - L^*) - 11.4(C^*)^{0.02}$	0.73

It is clear that Ou's model has relatively performed very well. Also the results of the multiple correlation technique show agreement amongst the models. For instance, heavy-light showed the greatest correlation between the three studies. It should be noted that the data gathered from these studies were based in different countries which made a variety of nationality amongst observers (Chinese, Japanese, British, Thai); the similarity which exists between these models suggests that these models are less dependent to cultural differences.

4.6.4 Critical analysis

It is important to note that in Ou's study, gender or cultural differences were not found. It has been concluded that there is no scientific evidence which explains the gender or cultural difference. This may not be true as the methodology of the psychophysical applied by Ou and others has limited the amount of variability between the participant's nationality and background. This is because the number of participants was not big enough to be able to draw a proper comparison between different cultural groups. In Ou's study for example, participants were required to answer orally according to definitions written in the dictionary which may cause bias in the core personal-connection between the colour and meaning. On the other hand answering back orally to the experimenter might cause judgmental-bias as the participant might find himself/herself trapped between what he /she thinks is right or wrong. Bearing in mind that there is no right or wrong and it is supposed to be based on how they feel about that colour. Some people might not be extrovert enough to be able to give their opinion orally but rather, it might be easier for them to write down or choose an option for their response. Another issue is that cultural differences may not show any significance with small number of people living all in the same country. To be clear, take the British and Chinese participants in Ou's experiment for example; both groups live and work in the same country which is UK. In psychology, it is a well established fact that psychological behaviour is strictly affected by the environment [170] and so if proper comparison has to be made amongst Chinese and British participants then these two groups must be in relevant countries, speak the native language and be questioned in the same language. Only in this way, can the difference be analysed as significant. Never the less, although in colour research controlled conditions of lighting and viewing distance are required to make the data and results scientifically reliable, the real world visualisation of colours are never the same and rarely are people presented with NCS colour samples in a complete dark environment, under D65 illuminant and a fixed grey background. It is time to move the colour experiments closer to the real world where people see colours from all sorts of dimensions, distances, illuminations and even materials. Finally; colour is a universal matter defined and named with all sorts of different ways around the globe making it a complex phenomenon which can only be visualised in a three dimensional space, precisely humans emotions are involved making it even more complex adding several more dimensions to it. So the perfect formulae should contain more than just a few variables that indicate colour attributes. For example; a model which is named as warm – cool not only should contain the colour attributes but also the cultural and gender variables too, so it would vary accordingly.

Therefore; the aim of this research is to shape a new type of approach (inspired by Moroney) for psychophysical experiments which can allow more variability in the data collection by rising the number of participants and varying the experimental environment. This is in order to form a real world situation for the experiment that initially leads to more reliable data which are not biased and can communicate a more comprehensible image of the existing relationships. Colours will be displayed on PC monitors all over the world with all sorts of random variation. This makes the experimental conditions very close to the real life where all the advertisements in the media are taking place (nowadays media takes advantage of the critical use of colours which is viewed mainly on pc or TV monitors). More people will be taking part from all over the world as the experiment will be online. Models will not be limited to colour attributes and will be dependent to all the effective variables.

5 Colour Semiotics Derived in the Laboratory

5.1 Introduction

Typically, colour experiments are carried out under controlled conditions in a laboratory with relatively few observers participating. This method is well known for its accuracy in colour production and measurements, since most of the affective variables in the viewing field can be kept constant (such as illumination, background colour, etc.). However, certain limitations (such as small numbers of participants and the durable simulation of the natural viewing conditions) apply to this experimental method. Later in the thesis, an alternative approach will be explored and implemented where the experiment will be carried out over the Internet in far less-well controlled conditions, but with large numbers of participants. In this chapter however, the conducting of a conventional laboratory-based experiment will be described. Part of the reason for this was to carry out the same experiment using two very different sets of experimental conditions so that a comparison is possible. In some sense, the experiment described in this chapter can be thought of as a control experiment and initially it would be used to test the hypothesis of whether the Global Online Colour Survey is effective or not. This laboratory experiment is based on Ous' experimental methodology [1] in which each participant rates each of 28 colours.

5.2 Methodology

The experiment is designed in the English language (since cultural effects are not being considered and individuals fluent in English are asked to participate). Each participant is required to rate each of 28 stimuli or colours (presented in random order) in terms of 10 bi-polar characteristics. Each stimulus is displayed on an extended grey background[63] (proximal field and background are the same colour) as illustrated in Figure 25[64]. A blank grey page is displayed between each colour in order to decrease any colour-adaptation effects. The stimuli were 4 cm × 4 cm and were viewed by observers who were seated approximately 50 cm from the screen. The experiment was conducted in a darkened room.

The monitor used for the experiment was a Lacey calibrated monitor located in the School of Design Colour Lab. The monitor was placed on a stable table and was switched on

63 RGB values of [190,190,190].

64 The Self-luminous displays (CRT and LCD monitors) are used in all of the experiments in this thesis and their measurements are carried out with the spectroradiometer.

approximately 15 to 30 minutes before the start of the experiment in order to warm up and approach the main calibrated settings. The monitor was placed around half a metre distant from the observer's seat. Therefore the relative colour values remained the same throughout the experiment.

In this study, each participant spent about 30 minutes taking part in the experiment. The rating or scaling technique used is an established method that is widely used in psychophysical experiments of this nature and is particularly widely used by the academic colour community. The method is known as categorical judgement (explained in 3.2.2) and methods for analysing data from such experiments have been known since the early part of the 20th Century.

Table 4: Bi-polar characteristics of interest in this study

Bi-polar characteristics	
Heavy	Light
Warm	Cool
Modern	Classic
Clean	Dirty
Active	Passive
Hard	Soft
Tense	Relaxed
Fresh	Stale
Masculine	Feminine
Like	Dislike

The experiment was implemented using a commercial service known as KeySurvey [171]. One of the reasons for this is that it allowed exactly the same presentation to be used both for this laboratory-based experiment and the subsequent internet-based experiment. Prior to the experiment each participant was asked about his or her gender, age, nationality and native language (Figure 24). The 28 colours that were used in the experiment are illustrated in

Table 5 along with the RGB values. The colours were chosen to approximately uniformly sample RGB space with some additional colours (for example, numbers 27 and 28) chosen. The 10 bi-polar characteristics are listed in Table 4; these were chosen to be consistent with the Ou experiment [1]. Participants are asked to rate each colour using a discrete scale (-2, -1, 0, 1, 2) as shown in Figure 25. The participants are instructed that the zero in the scale corresponds to the colour that have neither of the two bi-polar characteristics or an equal amount of each.

A total of 20 participants, fluent in the English language and with a basic knowledge about colour took part in the experiment. The participants successfully passed the Ishihara colour-blind test[65] before taking part[66]. The colour stimuli were specified in RGB values but, of course, it is necessary to know their CIE specifications. In this study, the CIE Yxy values were measured with a spectroradiometer (CS-100A, Luminance and colour meter) and are listed in Table 6 .The measured Yxy values were converted to LCH values in MATLAB using standard methods [11].

---

[65] This test is used to identify people with colour blindness. Participants are offered different coloured images in which numbers are suppose to be identified. If the participant is not successful in finding the numbers then he/she is colour blind.
[66] Only one participant from amongst the 21 participants was colour blind and his data was discarded.

Table 5: Table of the RGB values for the selected 28 colours (number 16 indicates white).

	Colour	R	G	B
1		255	255	0
2		128	0	0
3		0	64	255
4		0	128	0
5		128	128	255
6		128	128	0
7		255	128	255
8		128	255	255
9		128	64	128
10		0	128	128
11		128	255	128
12		255	0	255
13		128	64	64
14		128	128	64
15		0	255	255
16		255	255	255
17		0	0	255
18		0	255	0
19		255	0	0
20		64	128	64
21		128	128	128
22		64	64	128
23		128	0	128
24		255	255	128
25		0	0	0
26		255	128	128
27		64	128	128
28		255	127	0

Table 6: Yxy values measured by the spectra-radiometer and converted to LCH values.

	Colour	Y	x	y	L	C	H
1		75.70	0.39	0.52	89.72	87.96	111.20
2		7.04	0.54	0.34	31.90	46.51	33.05
3		9.78	0.15	0.08	37.44	122.04	301.50
4		23.50	0.29	0.57	55.58	74.78	139.60
5		37.00	0.22	0.20	67.28	56.87	286.07
6		29.10	0.39	0.51	60.87	59.37	113.03
7		47.20	0.28	0.23	74.32	55.10	309.30
8		75.10	0.24	0.31	89.44	35.28	216.29
9		18.10	0.28	0.23	49.62	36.67	308.55
10		26.50	0.28	0.31	58.51	11.18	238.13
11		70.10	0.29	0.43	87.05	53.79	152.02
12		26.10	0.27	0.15	58.13	105.86	318.91
13		16.50	0.35	0.33	47.62	12.22	13.62
14		30.40	0.34	0.41	62.00	27.99	126.18
15		69.80	0.21	0.30	86.90	49.65	210.63
16		85.50	0.28	0.31	94.10	15.17	240.12
17		4.71	0.16	0.09	25.89	79.30	299.25
18		62.80	0.30	0.59	83.34	109.60	137.85
19		17.80	0.59	0.34	49.25	77.28	38.27
20		27.40	0.29	0.42	59.34	38.18	151.99
21		32.90	0.28	0.32	64.08	11.00	225.93
22		15.10	0.23	0.22	45.77	37.33	284.42
23		10.40	0.27	0.16	38.55	71.41	318.27
24		81.30	0.34	0.41	92.27	39.54	124.38
25		1.83	0.28	0.32	14.57	4.88	216.27
26		43.40	0.36	0.32	71.83	20.65	11.54
27		29.20	0.24	0.31	60.96	24.57	211.48
28		42.70	0.46	0.45	71.35	66.39	80.49

.

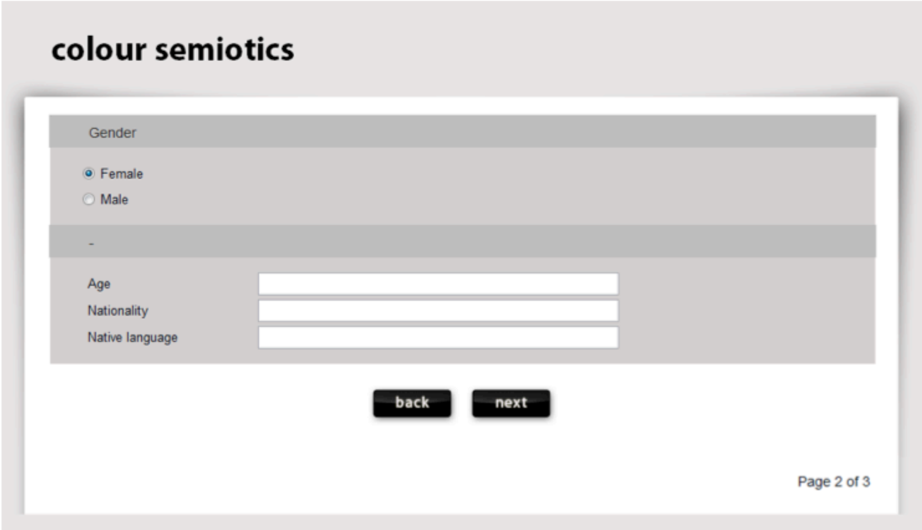


Figure 24: First page where the participant is asked about gender, age, nationality and native language.

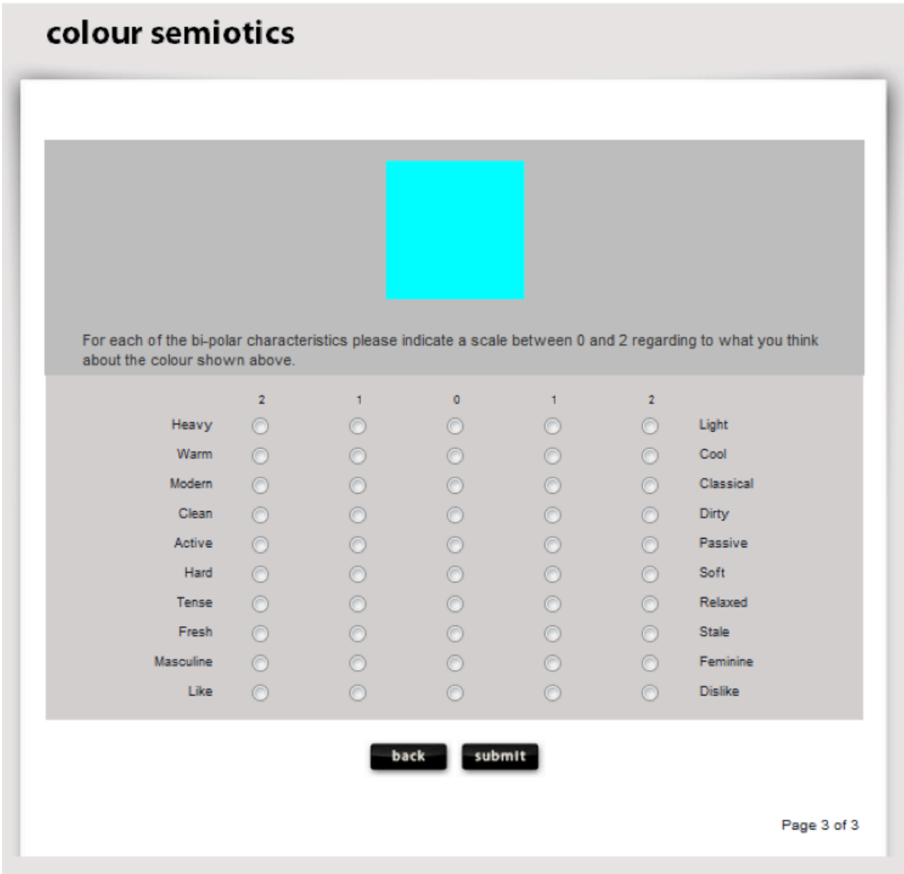


Figure 25: Second page where the participants have to rate the bipolar characteristics according to the colour displayed.

5.3 Data analysis

The data collected from the experiment is analysed in this section, with the aim of defining relationships between colours and bi-polar characteristics. The degree of agreement between participants is explored before the development of the final models. The Kappa statistic is one of the measures of agreement that has been proposed specifically for the categorical variables [172-177]. Initially, factor analysis is applied in order to find the affective characteristics, and finally model fitting takes place.

5.3.1 Participant agreement test

In order to determine the kappa statistic K; the coefficient of agreement for nominally scaled data, it is first required to produce a table for each bi-polar characteristic containing frequency values for all 28 colours. Taking heavy - light as an example; columns indicate the five different categories of -2, -1 , 0 , 1 and 2 and the rows contain the frequency responses for the heavy - light values for each of the 28 colours. According to Table 8, it is evident that the only bi-polar characteristic which shows significant difference of rating amongst the 28 colour range is like-dislike. As the derived z value (0.73) is less than the z value for significance level of $\alpha = 0.01$ which is 2.32. This means that by only using 20 observers, the portion of significant differences amongst the ratings is very low. On the other hand it shows that the like-dislike characteristic stands out from the other 9 which means that there could be certain facts about this word pair which needs to be explored, bearing in mind that this was the last word-pair in the survey[69][169]. In other words it can be concluded that for this certain type of experiment where limited number of response has been gathered, the only bi-polar characteristic that shows disagreement across the 28-colour range is like-dislike. By linking this to Ou's study, it can be said that it is caused by the cultural (or possibly gender) differences which exists between the respondents. The other

[69] As a short comparison with Ou's results; he too found that the correlation coefficient of like-dislike was to be very different amongst British and Chinese remarking the cultural effect. Since, the data collected from this study involved a diverse mixture of all cultures (and not very specific as Ou's), the study of cultural differences cannot be investigated in this part and will be explored using the lab-based experiment data collection.

[71] Note that, in the future, this programme can be coded as an interactive graph, presenting updated data collected from the Global Online Survey source.

9 bi-polar characteristics show agreements between observers, concluding that participant’s response are unlikely to be based on chance.

Table 7: Table of frequency for the Heavy - light characteristic categorised by the chosen scales for 28 colours

	Colour	-2	-1	0	1	2
1		9	3	3	4	1
2		0	0	1	6	13
3		1	2	4	5	8
4		0	4	4	7	5
5		5	9	3	3	0
6		2	1	7	6	4
7		9	10	0	1	0
8		17	1	1	1	0
9		0	1	3	9	7
10		0	8	4	7	1
11		11	8	0	1	0
12		2	7	2	6	3
13		0	3	3	6	8
14		2	3	3	7	5
15		12	7	0	0	1
16		15	4	0	1	0
17		0	0	2	7	11
18		10	6	2	1	1
19		0	3	2	7	8
20		0	3	5	8	4
21		1	5	3	8	3
22		2	2	3	5	8
23		0	3	0	9	8
24		12	7	1	0	0
25		1	0	1	2	16
26		6	6	4	3	1
27		4	6	3	6	1
28		4	6	3	6	1

Table 8: Table of the Kappa test values

	Kappa value	Var(Kappa)	Z
Active-passive	0.12	9.01e-05	12.73
Clean-dirty	0.11	8.90e-05	11.25
Fresh-stale	0.08	8.92e-05	8.38
Hard-soft	0.09	9.51e-05	9.18
Heavy - light	0.16	6.21e-05	20.27
Like-dislike	0.01	0.000193	0.73
Masculine-feminine	0.13	0.00021	9.01
Modern-classic	0.04	8.50e-05	4.72
Tense-relaxed	0.07	0.00011	7.04
Warm-cool	0.11	5.53e-05	14.18

5.3.2 Observer accuracy

Observer accuracy metric is used to understand the level of agreement between the participants for each of the ten bi-polar characteristics. It is important to derive fairly high values of observer accuracy since it can in-directly suggest possible correlations that can be modelled. This metric is used in Ou's research using the following equation:

Equation 66

$$Accuracy = \frac{\sum(1 - w_i)}{N}$$

Where N is the number of colours involved in the experiment and w_i is the proportion of observers who's decision were different from the majority. The observer accuracy values for both this experiment and Ou's are displayed in . According to the nature of the observer accuracy formulae, increase in the number of colours results in lower levels of observer accuracy. This could be one of the reasons behind lower values of observer accuracy for this experiment as 28 colours were used.

However, overall it seems that observer accuracy of both experiments is reasonably high and that there is fair agreement between the observers for each of the bi-polar characteristics. Heavy-light, warm-cool, hard-soft, active-passive and masculine feminine seem to have the highest agreement within the experiment. This agreement can conclude significant correlations that would be modelled further in this section.

Table 9: Table of observer accuracy values for the 10 characteristics.

	Experiment	Ou's experiment
HL	0.62	0.79
WC	0.6	0.85
MC	0.57	0.77
CD	0.61	0.7
AP	0.62	0.68
HS	0.59	0.65
TR	0.6	0.77
FS	0.58	0.69
MF	0.64	0.68
Total number of participant	20	30
Total number of colours	28	20

5.3.3 Categorical-judgment

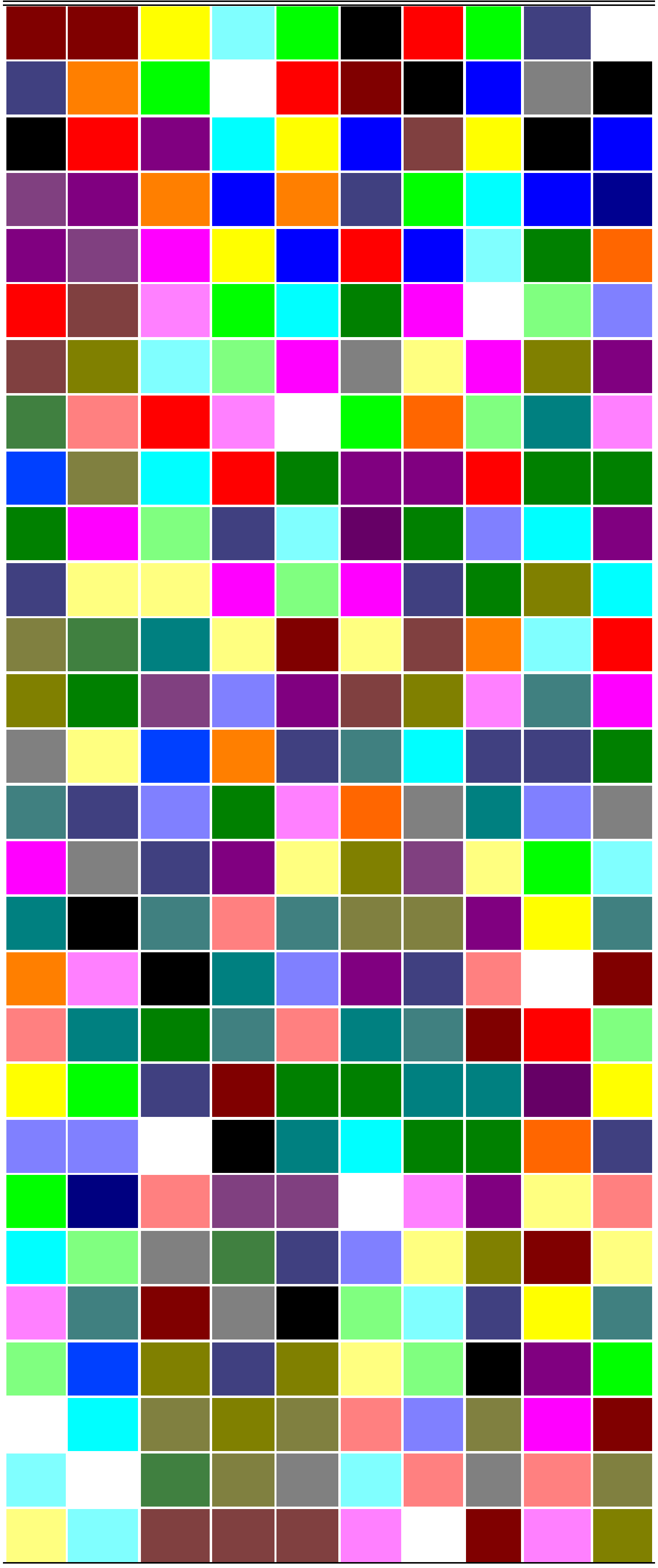
The categorical-judgment method can be described as an algorithm which takes the matrix of frequencies as the input and eventually returns the relative z scores (see 0 for the coded algorithm in MATLAB). The input matrix for this study is the frequency values of each of the bi-polar characteristics with each column defining the categories (-2, -1, 0, 1 and 2) and 28 rows standing for the 28 colours. The categorical-judgment algorithm leads to two sets of results; one is the ranks which are available for the variables defined in the rows and the other is a vector of z scores. In this section, both of these outputs are analysed.

5.3.3.1 Ranks

By applying the categorical-judgment algorithm to the data, ranks can be generated as an output. Thus, 28 colours are ranked in respect of each bi-polar characteristic. A graphical illustration of the ranks is displayed in the following Table. Lower rank value indicates the first word from the pair (e.g. heavy from heavy - light) and higher rank indicates the second word from the pair (e.g. light). For example; colour number 1 (yellow) has been ranked the 20th colour whilst colour number 2 (red) is ranked first for heavy - light which means that colour number 2 is ranked to be the heaviest amongst the 28 colours and colour number 1 (yellow) appears to be less heavy (to me more light) and stands at the 20th position, in this range colour number 24 tends to be the lightest.

Table 10: Ranking of colours for each of the bi-polar characteristics (for example, in the heavy-light column the heaviest colours are shown at the top and the lightest at the bottom).

HL	WC	MC	CD	AP	HS	TR	FS	MF	LD



5.3.3.1.1 Kendall's coefficient of concordance W

At this point, it is very useful to understand if there is a certain concordance or in other words to measure the agreement between the bi-polar characteristics. The ranked data need to be dealt with using a non-parametric method. According to the null hypothesis of this test the 10 sets of ratings for the bi-polar characteristics is unrelated or independent. Hereby according to the small significance value of 0.001, it is concluded that the null hypothesis of independence is rejected and that there is a degree of relation (almost 0.21) amongst the colour rankings of the 10 bi-polar characteristics.

Table 11: Kendall's test of concordance

	Kendall's W	X^2	Degree of freedom	Probability of significance
Values	0.21	57.83	27	0.001

The algorithm has been programmed in MATLAB and the results suggest that there is significant agreement amongst the 10 bi-polar characteristics (ratings of level 0.21). This indicates that some or all of these characteristics could be correlated which will be further investigated by the study of the correlation matrix and factorial analysis.

5.3.3.2 Z scores

Rank values were analysed in the previous section. However, there are some disadvantages of using the rank values since they are not fully informative about the interval nature of the collected data. This fact can be explained by an example of five runners named as A, B, C, D and E, whom reach the ending line in an order where A will be ranked as first, B as second, C as third and so on. However, A has a time lap of 10 minutes, whilst B has 9.85, C and D both have 9.5 and E has a time lap of 9 minutes. The ranks only inform about the order, rather than the intervals between the runners. Likewise, measuring the average of ranks may be meaningless. One of the main aims of this study, by using the Likert's scale and performing the test with a categorical-judgment on the colours, is to obtain interval data. So for each bi-polar characteristic the colour's rates are identified throughout the scales. Or in other words, each colour's individual scale is identified for a certain bi-polar characteristic. In this section, the Z scores derived from the categorical-judgment method are analysed. It is also possible to calculate the categories for each of the Z scores.

5.3.3.2.1 Results

Table 13 contains the Z scores derived from the categorical-judgment method for each of the 10 bi-polar characteristics, across the 28 colours. For example, the first row of this table holds the Z scores for each of the ten bi-polar characteristics for the colour yellow. Interestingly, this colour seems to have a neutral scale for active–passive, hard–soft and like–dislike. It is rated 1 for being warm, modern, clean, tense and fresh whilst having the same scale for being light and feminine. A graphical presentation of the colour rates for each of the bi-polar characteristics is displayed in Figure 26. The figures are a visual presentation of the Z score ranks in which the inner circle indicates rates equal to 1 and the outer circle indicates Z score ranks of 2. Note that the colours, which are not shown in the two circles, are rated as zero. Also, like-dislike is not included in the figure since the Z scores related to this bi-polar characteristic are not categorised.

Table 12: Table of Z score ranks.

	Colour	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
1		20	11	1	5	3	12	7	3	24	20
2		1	1	24	20	12	2	3	19	23	18
3		9	25	14	4	5	3	5	2	4	3
4		10	13	19	15	9	6	10	11	9	9
5		21	21	15	13	18	23	26	10	15	6
6		13	7	25	26	25	17	13	23	11	28
7		24	18	6	8	15	28	22	13	28	8
8		27	28	7	1	10	27	24	5	12	16
9		4	5	13	22	22	18	16	22	20	7
10		15	19	12	18	17	19	19	15	13	17
11		25	23	10	7	11	24	25	8	18	19
12		16	10	5	11	7	11	6	7	26	13
13		7	6	28	28	28	13	12	28	6	26
14		12	9	26	27	26	16	17	26	7	27
15		23	26	9	3	6	21	14	3	10	11
16		26	27	21	2	8	22	28	6	17	1
17		2	22	16	10	14	4	11	14	1	4
18		22	20	2	6	1	8	4	1	16	25
19		6	3	8	9	2	5	1	9	19	12
20		8	12	27	23	20	20	21	21	5	14
21		14	16	23	24	27	7	15	27	2	15
22		11	15	20	25	23	10	18	24	14	21
23		5	4	3	16	13	9	9	17	25	10
24		28	14	11	12	16	25	23	16	22	23
25		3	17	18	21	24	1	2	25	3	2
26		19	8	22	17	19	26	27	18	27	22
27		17	24	17	19	21	13	20	20	8	24
28		18	2	4	14	4	15	8	12	21	5

Table 13: Table of Z scores.

	Colour	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
1		0.68	1.47	1.86	2.39	2.50	1.57	2.06	2.38	0.78	1.62
2		2.57	2.52	0.79	1.33	1.87	2.63	2.46	1.37	0.91	1.69
3		1.75	0.43	1.26	2.39	2.29	2.49	2.30	2.46	2.46	2.32
4		1.75	1.31	1.03	1.96	1.95	1.82	1.78	2.04	2.07	1.95
5		0.51	0.67	1.23	2.11	1.32	0.83	0.96	2.07	1.51	2.11
6		1.48	1.82	0.60	0.79	0.72	1.51	1.68	1.00	1.74	1.18
7		0.04	0.88	1.73	2.27	1.66	0.53	1.06	1.92	0.32	2.00
8		-0.12	-0.24	1.67	2.71	1.90	0.69	1.02	2.33	1.71	1.77
9		2.09	1.90	1.30	1.22	0.93	1.49	1.52	1.01	1.15	2.05
10		1.33	0.69	1.33	1.50	1.37	1.20	1.42	1.63	1.67	1.70
11		-0.01	0.55	1.46	2.28	1.89	0.76	1.01	2.17	1.24	1.63
12		1.23	1.57	1.75	2.19	2.06	1.65	2.26	2.19	0.54	1.80
13		1.93	1.88	0.49	0.59	0.35	1.56	1.70	0.52	2.15	1.29
14		1.49	1.76	0.55	0.68	0.53	1.53	1.49	0.64	2.14	1.21
15		0.12	0.29	1.60	2.48	2.14	1.03	1.63	2.38	1.76	1.90
16		-0.12	0.25	0.99	2.64	1.97	0.86	0.83	2.23	1.28	2.37
17		2.44	0.66	1.18	2.23	1.76	2.26	1.70	1.80	2.67	2.18
18		0.41	0.67	1.85	2.32	2.80	1.77	2.33	2.65	1.48	1.32
19		1.96	1.96	1.64	2.24	2.77	2.22	2.77	2.12	1.21	1.85
20		1.76	1.34	0.50	0.91	1.02	1.12	1.10	1.07	2.19	1.79
21		1.43	0.99	0.88	0.85	0.44	1.79	1.60	0.61	2.55	1.78
22		1.64	1.10	1.02	0.83	0.87	1.69	1.44	0.83	1.66	1.60
23		2.04	1.96	1.79	1.82	1.84	1.71	2.01	1.50	0.71	1.90
24		-0.19	1.27	1.35	2.16	1.46	0.73	1.02	1.61	0.91	1.55
25		2.28	0.90	1.06	1.32	0.83	2.87	2.51	0.75	2.53	2.35
26		0.72	1.79	0.91	1.61	1.10	0.72	0.85	1.39	0.49	1.55
27		0.68	1.47	1.86	2.39	2.50	1.57	2.06	2.38	0.78	1.62
28		2.57	2.52	0.79	1.33	1.87	2.63	2.46	1.37	0.91	1.69

Table 14: Table of categories defined according to the Z scores.

	Colour	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
1		-1	1	1	1	0	0	1	1	-1	0
2		2	2	0	0	1	1	0	0	0	0
3		1	-1	1	1	1	1	1	0	1	1
4		1	0	0	0	0	0	1	0	0	1
5		-1	-1	1	0	-1	-1	1	0	0	1
6		1	1	-1	-1	0	0	-1	-1	0	0
7		-1	-1	1	0	-1	-1	0	1	-1	1
8		-2	-2	1	0	-1	-1	1	1	0	0
9		1	1	0	-1	0	0	-1	0	0	1
10		0	-1	0	0	0	0	0	0	0	0
11		-2	-1	1	0	-1	-1	1	0	0	0
12		0	1	1	1	0	1	1	1	-1	0
13		1	1	-1	-1	0	0	-1	-1	0	0
14		1	1	-1	-1	0	0	-1	-1	0	0
15		-1	-1	1	1	-1	0	1	1	0	0
16		-2	-1	1	1	-1	-1	1	0	0	1
17		2	-1	1	0	1	0	0	0	1	1
18		-1	-1	1	1	0	1	1	1	0	0
19		1	1	1	1	1	1	1	1	0	0
20		1	0	-1	-1	-1	-1	0	-1	0	0
21		1	0	-1	-1	0	0	-1	0	1	0
22		1	0	-1	-1	0	0	-1	0	0	0
23		1	1	0	0	0	0	0	1	-1	0
24		-2	0	1	0	-1	-1	0	0	0	0
25		1	-1	0	-1	1	1	-1	0	1	1
26		-1	1	0	-1	-1	-1	0	0	-1	0
27		-1	-1	0	-1	0	-1	0	0	0	0
28		-1	1	0	1	0	0	0	1	0	1

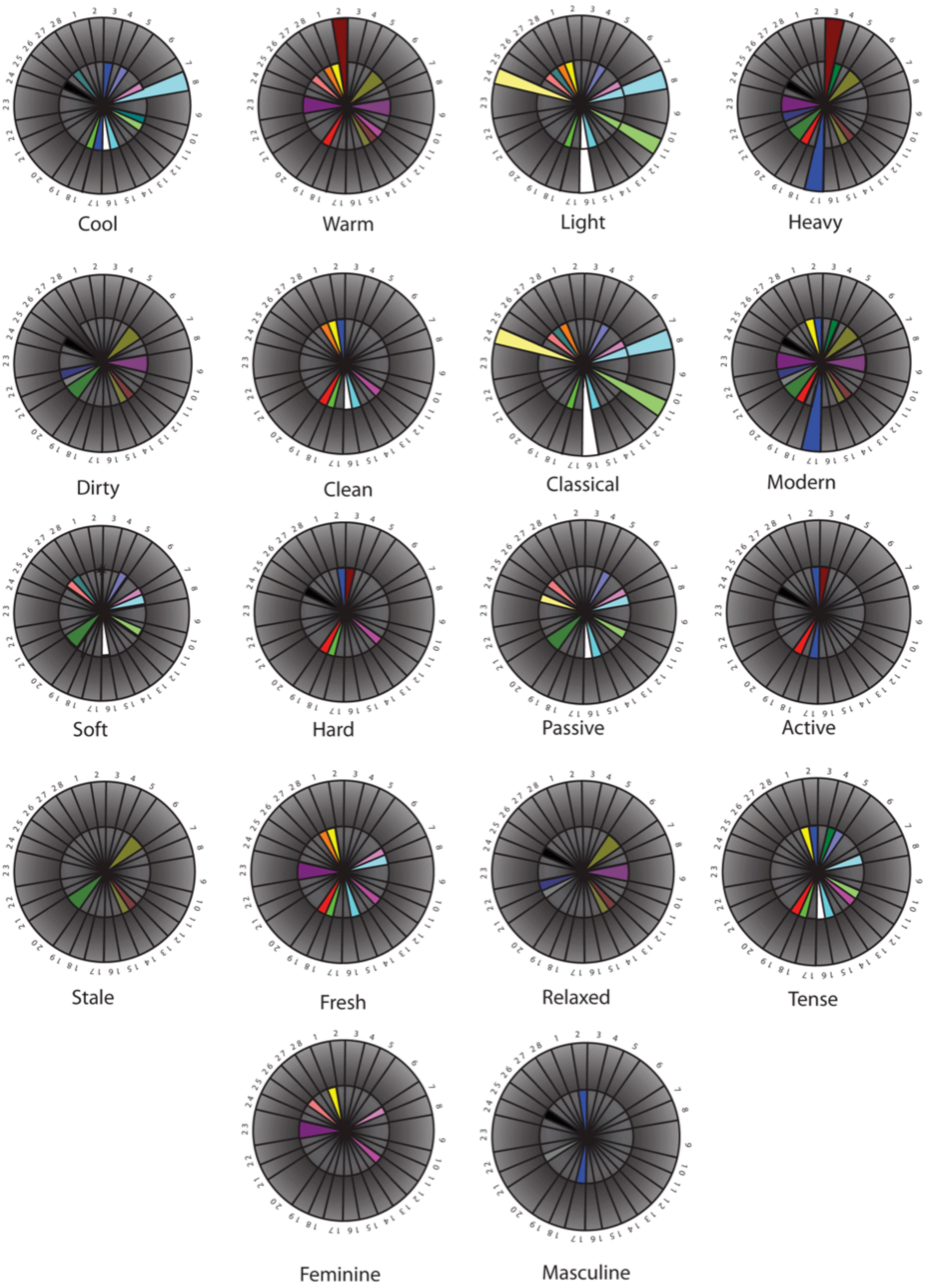


Figure 26: Illustration of the colours associated with the characteristics (inner circle indicates a rate of 1 and outer circle indicates 2).

Additionally, it is beneficial to report the Z scores by a user-friendly tool, rather than the usual tables. This aids the effective communication of the results of this experiment, and therefore undertakes the main motive of this thesis, which is to bridge the gap between designers and scientists by using understandable and simplistic language. Therefore, an interactive interface is programmed in MATLAB and displayed in Figure 27[71].

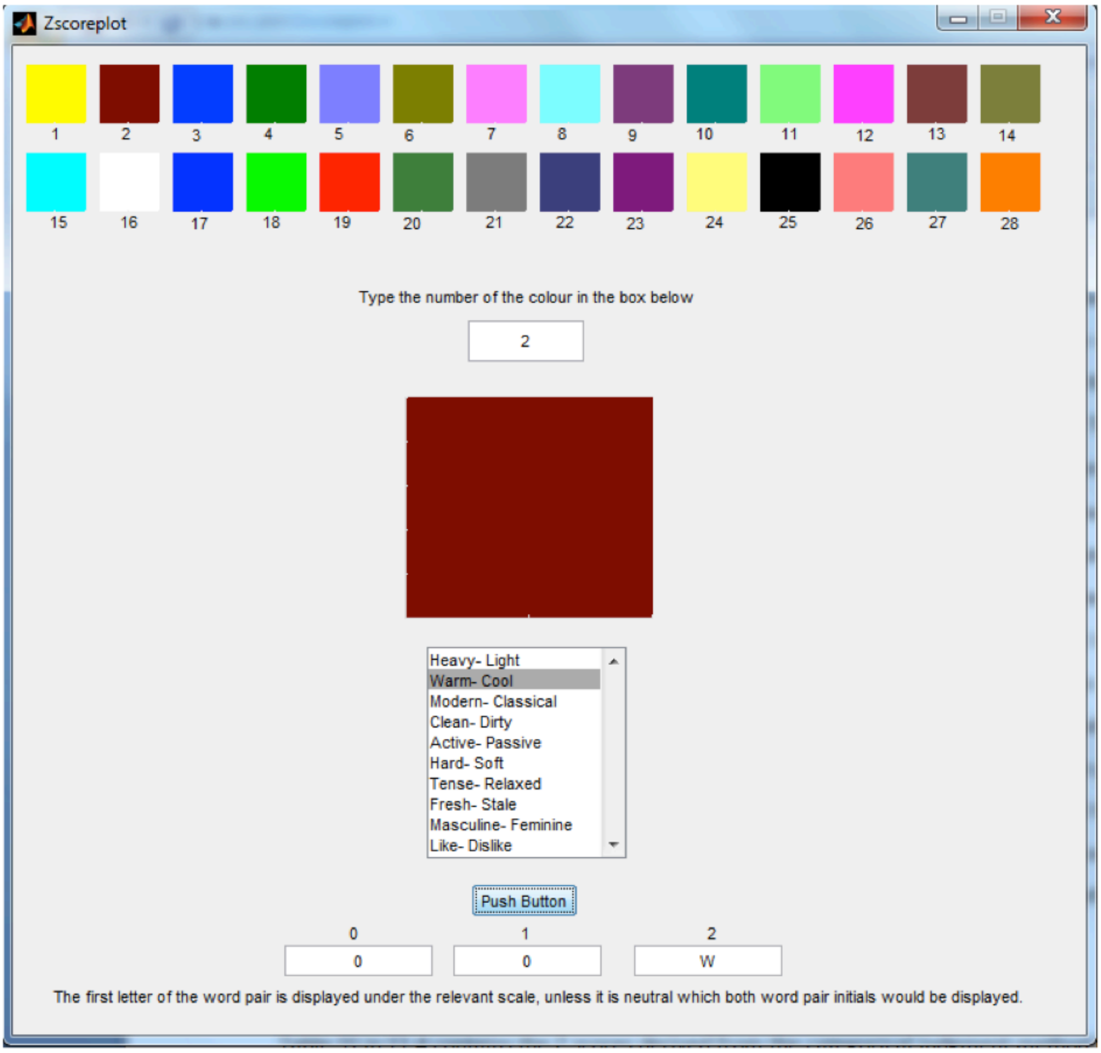


Figure 27: Interface of the tool illustrating the ranking of the Z scores for each of the bi-polar characteristics regarding the chosen colours.

As displayed in Figure 27, the user would simply type the number of the desirable colour and choose a bi-polar characteristic from the list. By hitting the button, the first letter of the bi-polar characteristic would appear under the relevant scale number 0, 1 or 2. The first

---

[71] Note that, in the future, this programme can be coded as an interactive graph, presenting updated data collected from the Global Online Survey source.

letter of the first word pair is used to illustrate the positive Z score, and the first letter of the second word pair is used to illustrate the negative Z score. Both initials would be used for the Z score equal to 0 (since they don't have precedence over each other).

5.3.3.3 Correlation Matrix for the bi-polar characteristics

The strength of the inter-correlations among the bi-polar characteristics is studied by Table 15 in which values more than 0.3 are considered to report a significant correlation [178]. As an example, in the first row of the correlation matrix illustrated; the highest correlation coefficient is 0.805, which describes the correlation of heavy - light and hard-soft characteristics. In other words, these two bi-polar characteristics have high positive relationship together pointing to the fact that higher value for heaviness correlates with higher level of hardness with the degree of 0.805. Conversely; lower level of heaviness (higher lightness) correlates with lower level of hardness (higher softness). Another example is the negative correlation coefficients of the clean-dirty and heavy-light bi-polar characteristic of -0.56 which informs of a reverse correlation, indicating that higher level of cleanness leads to lower level of heaviness and higher level of dirtiness leads to lower level lightness.

Table 15: Correlation Matrix of the bi-polar characteristics.

	HL	**WC**	**MC**	**CD**	**AP**	**HS**	**TR**	**FS**	**MF**	**LD**
HL	1.000	.537	-.404	-.562	-.270	.805	.587	-.515	.380	.089
WC	.537	1.000	-.204	-.480	-.139	.307	.409	-.411	-.363	-.270
MC	-.404	-.204	1.000	.762	.762	-.130	.243	.732	-.517	.272
CD	-.562	-.480	.762	1.000	.844	-.235	.008	.939	-.351	.467
AP	-.270	-.139	.762	.844	1.000	.078	.398	.910	-.382	.310
HS	.805	.307	-.130	-.235	.078	1.000	.848	-.196	.419	.188
TR	.587	.409	.243	.008	.398	.848	1.000	.102	.074	.107
FS	-.515	-.411	.732	.939	.910	-.196	.102	1.000	-.339	.329
MF	.380	-.363	-.517	-.351	-.382	.419	.074	-.339	1.000	.070
LD	.089	-.270	.272	.467	.310	.188	.107	.329	.070	1.000

A graphical illustration of the correlation matrix may be more effective in terms of communicating the links between the bi-polar characteristics and easier interpretation. Thus a graphical illustration is programmed in MATALB and displayed in Figure 28.

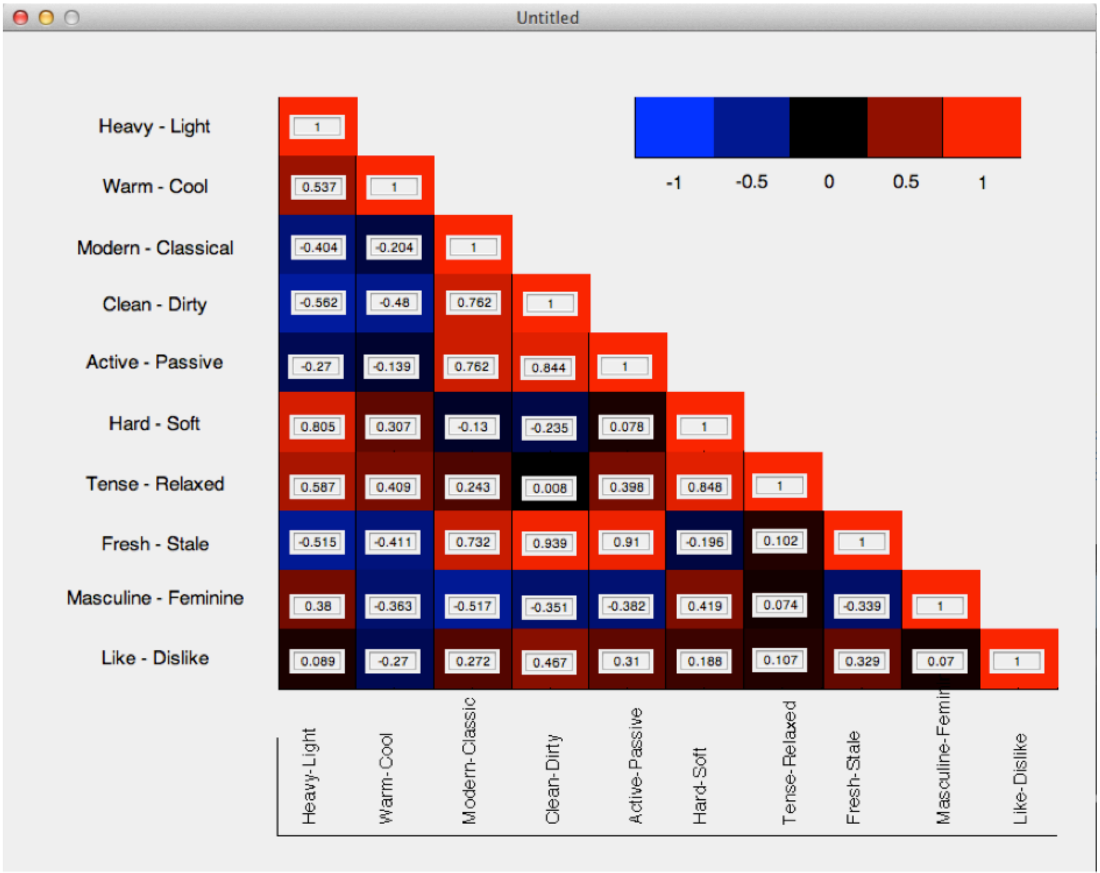


Figure 28: Colour chart for the correlation matrix (coded in MATLAB).

It is evident that fresh-stale is highly correlated with modern-classical, clean-dirty and active-passive. Also active–passive is highly correlated with modern–classical and clean–dirty. Clean- dirty is negatively correlated with warm-cool and heavy–light. Also masculine–feminine and modern–classic are negatively correlated. Note that correlation matrix is used to show the inter-correlation between each bi-polar characteristic, reporting the extent to which bi-polar characteristics are like each other. Correlation between two variables may not imply causation, since there may be a third variable; causing the affect that is based on the correlations between two or more of the variables. The existence of this third variable can be investigated using the factor analysis. In other words, the correlation matrix indicates the linear relationship between two variables, whilst factor analysis identifies variables that all relate to a single factor by exploring the interactions between them [179]. According to the correlation values (mostly higher than 0.3) there is likelihood that a factor structure underlies the variable. Therefore it will be further explored in next section.

5.3.3.4 Factor analysis

Regarding to section 3.3.1 of the literature review, one of the applications of factor analysis is the principal component method of extraction. Factors that explain an inter-correlation among the variables are identified through principal component analysis. Note that the terminology of the word factor is used differently in ANOVA, which indicates an independent variable rather than a dimension, which underlies several variables. This shows that in principal component analysis we are not interested in the individual variables, but the pattern of correlations [179]. The Z scores and the correlation matrix derived from the categorical-judgment technique have been used as the input of the principal component analysis.

5.3.3.4.1 Principal components methods of extraction

It is crucial to test whether the use of principal component analysis is appropriate on the collected data or not. Kaiser-Meyer-Olkin Measure of Sampling Adequacy [180, 181] and Bartlett's Test of Sphericity are used to test this hypothesis [182]. Table 16 holds the KMO value of 0.61 and a Bartlett's significance level less than 0.05 indicating that the usage of principal component analysis for this data is appropriate [178]. Therefore, it is possible to proceed to the next stage of the analysis, which is to derive the number of crucial factors.

The number of components is recognised by the Kaiser's criterion. It indicates that the components with Eigen values of 1 or more are crucial and have to be taken as individual components (Table 16). The Eigen values of the first three components are determined to be more than 1 (4.38, 2.675 and 1.577), which altogether explain a total of 80.895 percent of the total variance.

Another source, which is used for identifying the number of components, is the Scree Plot. It is a graphical illustration of the monotonically descending Eigen values. It is intended to aid the identification of the "trivial" components [183]. For example in Figure 29; components 1, 2 and 3 seem trivial, compared to the other remaining components. This suggests that initially, three components may best describe the data and be sufficient.

Table 16: KMO and Bartlett's Appropriateness test for Principal component analysis.

Test	Values
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.611
Bartlett's Test of Sphericity significance	0.00

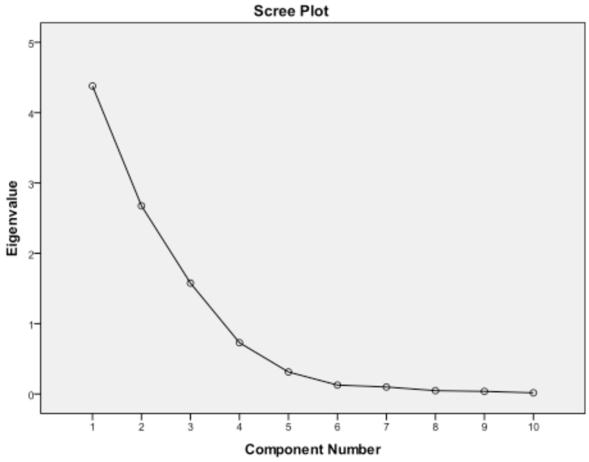


Figure 29: Scree plot for determination of the number of crucial components.

Table 17: Component Matrix of the three factors.

	Component		
	1	2	3
HL	-.699	.636	.050
WC	-.468	.355	-.745
MC	.826	.285	-.190
CD	.959	.142	.141
AP	.818	.492	-.090
HS	-.407	.863	.199
TR	-.093	.943	-.116
FS	.933	.200	.066
MF	-.486	.069	.798
LD	.348	.353	.510

The next step is to find out which bi-polar characteristic is to be categorised in the three components. This is identified by the use of component matrix illustrated in Table 17. In other words, this table contains the correlation values of each variable with each factor[72]. The contribution of the variables on each component is called loading[73].

The Varimax rotation method is used to interpret the three components which assumes that the factors are independent and decreases the number of variables that have high loadings on each factor [178][74]. Table 18 illustrates the rotated component matrix. Take the first row for example; the heavy-light loading on component 1 is -0.418, component 2 is 0.849 and component 3 is 0. Thus, component 2 with the highest loading of heavy-light will indicate the principal component containing other bi-polar characteristics which are alike.

Table 18: Pure bi-polar characteristic loadings on three components [Rotated component matrix], using the Varimax rotation method with Kaiser Normalisation (Rotation converged in 4 iterations).

	Component		
	1	2	3
HL	-.418	.849	.000
WC	-.239	.490	-.777
MC	.886	-.042	-.113
CD	.928	-.218	.225
AP	.946	.155	-.011
HS	-.077	.955	.178
TR	.270	.909	-.107
FS	.932	-.155	.149
MF	-.493	.256	.755
LD	.408	.208	.544

[72] . Note that the results of the Common Factor analysis are very similar to the loadings. However, the results depend on the data, number of variables and their communalities and may show significant differences for other data.

[73] SPSS uses the Kaiser criterion, which takes the components with values above 1, in to account.

[74] The best scenario is the so-called 'simple structure' in which the loading of each, are only on one of the components. Which, unfortunately, does not appear here, but still leaves the results in a reasonable condition to be interpreted.

The three-dimensional plot of the three components is displayed in Figure 30. It is a good visual presentation of the bi-polar characteristics and their allocations according to the three components.

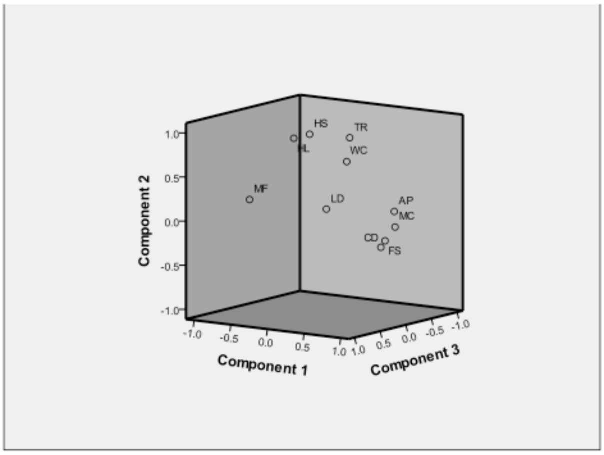


Figure 30: Three-dimensional plot of the three components and the placement of the ten bi-polar characteristics in rotated space.

The component score coefficient matrix illustrates values that can be used as sets of weights for the computation of the factor scores. Initially, they can affectively be applied for interpretation. For example, the contribution of the bi-polar characteristic heavy-light to component 1 is equal to -0.0603, component 2 is 0.28 and component 3 is 0.022 (factorised as component 1).

Table 20 contains the bi-polar characteristics according to their loadings on each component. The bi-polar characteristics loaded in component 1 and 2 have positive correlation with each other. Taking component 1 for example, if a colour is rated as heavy, it will also be modern, clean, active and fresh. However, in component 3, if a colour is cool then it is also masculine and liked. Conversely if it is warm, then it is feminine and disliked. Therefore the term warm-cool is replaced by cool-warm to indicate the reverse relationship.

Table 19: Component score coefficient matrix of the three components (using the Varimax rotation method with Kaiser Normalisation, rotation converged in 4 iterations).

	Component		
	1	2	3
HL	-.063	.280	.022
WC	-.010	.155	-.478
MC	.224	.028	-.102
CD	.215	-.030	.109
AP	.245	.101	-.038
HS	.021	.336	.123
TR	.116	.334	-.069
FS	.221	-.008	.061
MF	-.137	.073	.495
LD	.094	.099	.332

Table 20: Three primary factors identified in this research (note warm-cool has a negative correlation).

Factors	Bi-polar characteristics
Component 1	Modern-classic, Clean-dirty, Active-passive, Fresh-stale
Component 2	Heavy-light, Hard-soft, Tense-relaxed,
Component 3	Cool-warm, Masculine-feminine, Like-dislike

5.3.3.4.2 Conclusion

The 10 bi-polar characteristics were subject to principal component analysis. The suitability of the principal component analysis has been tested using the Kaiser-Mayer-Oklin and Bartlett's test, which results, revealed the significance of the test. Eventually; principal component analysis revealed 3 components with Eigen values higher than 1. The three components were developed by the Varimax rotation technique and the results revealed three components with strong bi-polar characteristic loadings named as evaluation, weight and quality. The results of the factor analysis are not very far from the findings of Osgood, Ou and etc.

5.3.4 Summary of data analysis

In this section, the collected data from the experiment has been analysed using the categorical-judgment technique resulting the ranks and Z scores on which factorial analysis has been carried out. Defining three principal components and also extraction of some bi-polar characteristics has been a subject of this study. The illustration of the Z scores has been made in an effective way (interactive tool programmed in MATLAB). However, the results derived from this section are to be used for modelling the relationship between colours and bi-polar characteristics. Section 6 is dedicated to modelling these relationships using the findings of this section.

5.4 Comparison with Ou's research

In terms of structure and methodology, this experiment has been designed to be very similar to Ou's research. However in this study, a step forward has been taken towards collecting more information through different scaling technique which is the Likert's categorical-judgment scaling. Also, in Ou's research 20 colours are chosen from the NCS Atlas and displayed in the lighting cabinet for observers. These colours were scattered across the CIELAB colour space covering a large range of hue, lightness and chroma (see Figure 31). The question and answers were taken place orally, and dictionary definitions of the word pairs were shown to each observer. In the Lab-based experiment, 28 colour samples were displayed in the monitor and were chosen according to the RGB values displayed. As a comparison; Figure 22 and Figure 31 display the scatter diagram of the colours in the L*a*b* space. With respect to the difference between the natures of the two colour spaces (NCS atlas and RGB values), the figures still show similar distribution.

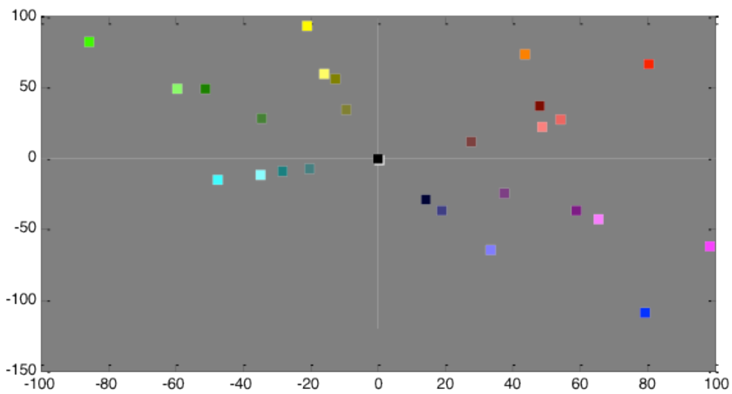


Figure 31: The 28 colour samples used in the Lab-based experiment, displayed in the CIELAB colour space.

Ou's experiment was carried out with 31 observers, which consisted of 14 British and 17 Chinese native speakers. This was organised in order to discover the effect of culture. However, the Lab-based experiment was carried out on 20 people from different sorts of backgrounds. In this research, it is believed that the effect of culture needs to be explored through mass data collection. Therefore, the effect of culture is to be investigated later in the online colour survey. Ou's goal was set to clarify the relationship between colour preference and colour emotion, and he also derived four significant models (heavy-light, warm-cool, active-passive and hard-soft). In this section the data gathered from the Lab-based experiment are compared with the values derived from inputting the colour values in to Ou's models.

5.4.1 Data analysis

Four sets of formula on heavy - light, warm-cool, hard soft and active-passive were found to be significant in Ou's research that would be used in this section. These formulae are based on the LCH values and therefore the 28 colour values used in the Lab-based experiment are placed into Ou's formulae resulting numerical values for heavy-light, warm-cool, active-passive and hard-soft. Ou used case V of Thurstone's law of ccomparative-judgment, announcing Z scores that indicated the extent to which different observers associated a colour sample with the first word of a bi-polar characteristic. So in other words, a large z score indicates high percentage of observers choosing the first word from the word pair. On the other hand, a positive z score indicated the frequency percentage for the first word and the negative for the second. However, in the Lab-based experiment, the scaling technique is based on ordinal sets, as each observer would not only associate a single word to a colour but to rate it too (scale it to either 1 or 2). Also, the user can choose the scale zero where none of the two words are suitable for the colour (see Figure 11). So the data collected in this survey are categorised as ordinal scales, well known as Likert's scale. The comparison between the two methods would be complex. To overcome this difficulty it has been decided to use method and derive Z scores.

5.4.2 Comparison of the principal components

Ou has used the factor analysis technique for finding the components, and has derived similar results. Apart from that, in Ou's work, like-dislike was excluded from the factors and warm-cool had been considered as a standalone factor, see Table 21. It points out that no matter if the data scaling technique is paired comparison (Ou's approach) or categorical-judgment (Likert's scale); the bi-polar characteristics would still be factorised in to three components with a slight difference. In comparison with Osgood's primary factors indicated in Table 22 it is evident that some of the variables have fallen into the same components such as heavy-light and hard-soft. However, Osgood divided the bi-polar characteristics according to their literal meanings and named them as evaluative, potency and activity [162] rather than using factorial analysis. The factor agreement which exists between Ou's data and this work (Table 20), also draws an indirect agreement with Kobayashi's three colour image scale (clear-greyish, soft-hard and warm-cool) [167] and Sato's three colour emotion categories (which is activity, potency and warm-cool) [160]. Note that initially Ou modelled the relationship between colours and each of the three components derived from the factor analysis, and also defined a three-dimensional colour space for colours based upon the three components.

Table 21: Three factors identified by Ou.

Factors	Bi-polar characteristics
Colour activity	Clean-dirty, Fresh-stale, Active-passive, Modern- classic
Colour weight	Heavy - light, Hard-soft, Masculine-feminine
Colour heat	Warm-cool

Table 22: Three primary factors identified by Osgood.

Factors	Bi-polar characteristics
Evaluative	Clean-dirty, Fresh-stale, Like- dislike
Potency	Heavy - light, Hard-soft, Masculine-feminine
Activity	Warm-cool, Modern-classical, Active-passive, Tense-relaxed

5.4.3 Modelling the relationship between colours and bi-polar characteristics

The Z scores derived in the previous section and the Yxy values measured from the colour measurements are used to model the relationships between colours and the bi-polar characteristics. Since the Z scores derived from the categorical scales indicate interval data they tolerate the application of regression models on data that were primarily collected with a discrete nature. In this section the approach to modelling the relationship between the bi-polar characteristics and colour, is taken by the consideration of Ou's models. In other words, in this section a preliminary step towards modelling the relationships would be taken by using the same format as Ou's models but with different coefficients. This aids the further understanding of the data behaviour as well as bridging the current modelling techniques to the more effective ones. So, this method is used to derive equations that are similarly formatted with Ou's equations and at the same time fit the data well. Several techniques are used in order to find the suitable coefficients which best match ou's form of equations that are explained in this section.

5.4.3.1 Methodology

The Yxy values of the 28 colours on the display monitor are measured using a spectrophotometer. The values are then turned in to LCH values using MATLAB conversion programme. Initially they are applied to Ou's formulae for heavy-light, warm-cool, hard-soft and active-passive. This data are then compared with the data gathered from the experiment. Figure 32 displays the scatter diagram for the 4 bi-polar characteristics that illustrates the relationship between the two sets of data. It is evident that all four have correlation; however heavy-light and active-passive seem to have a closer relationship. Table 23 contains the R^2 values of these relationships. Indicating that 85 percent of the data gathered from the experiment for heavy-light is in agreement with Ou's formulae. However, only 52 percent of the data collected for active-passive is in correlation with Ou's formulae. This means that the data gathered from the experiment are not constantly correlated with Ou's formulae.

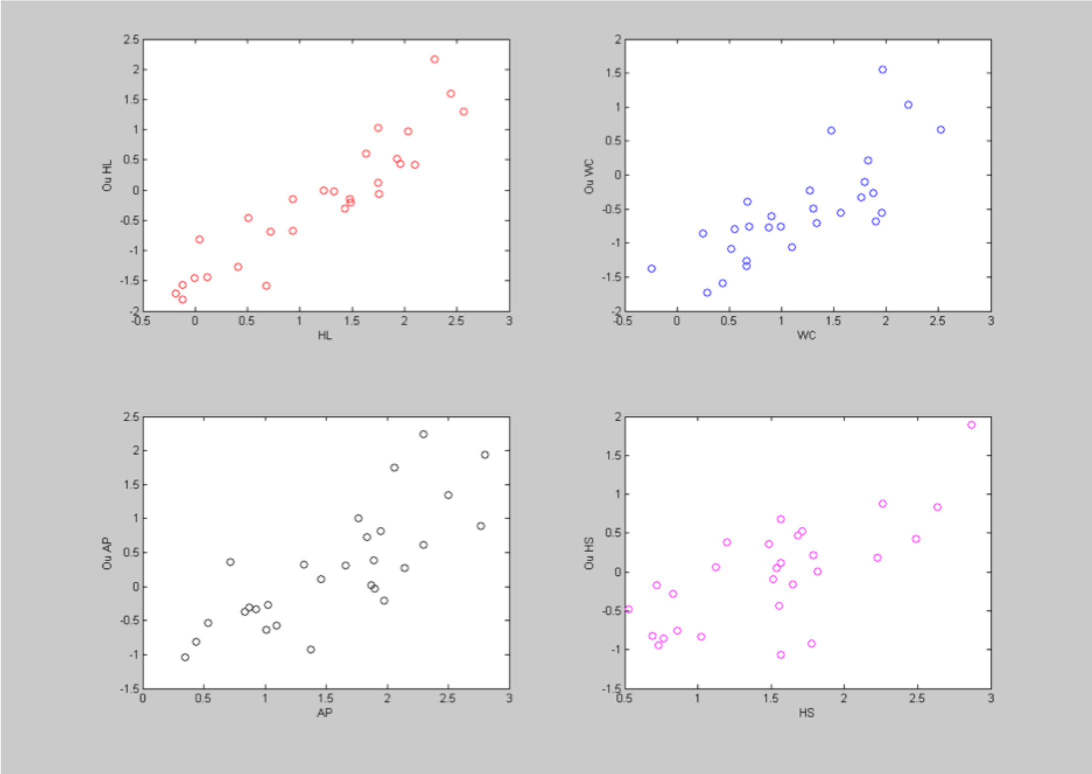


Figure 32: The scatter diagram for the relationship between heavy-light (HL) data collected from the experiment and heavy-light values derived from Ou's formulae (Ou HL). (Also for warm-cool (WC), active-passive (AP), hard-soft (HS)).

Table 23: R^2 values for the relationship between data collected from the experiment and values derived from Ou's formulae.

Ou's equations	**Model R^2**	$R^2{}_{data,model}$
$\boldsymbol{HL = -2.1 + 0.05\,(100 - L^*)}$	0.76	0.85
$\boldsymbol{WC = -0.5 + 0.02(C^*)^{1.07}\cos(\mathrm{H} - 50°)}$	0.74	0.59
$\boldsymbol{HS = 11.1 + 0.03(100 - L^*) - 11.4(C^*)^{0.02}}$	0.73	0.61
$\boldsymbol{AP = -1.1 + 0.03[(\Delta C^*)^2 + (\Delta L/1.5)^2]^{1/2}}$	0.75	0.52

5.4.3.2 Results

Approximated models, which are formed according to Ou's equations, are displayed in Table 24. Note that in this section, only the major coefficients are to be explored. Model 1 is derived from the simple matrix model fitting technique. The Approximation technique applied, is basically a stepwise method that slightly changes the coefficients of a regular regression model in each step. The R^2 is used as a basis of judgment in each step so that the coefficients are chosen so that they perform similar to Ou's models, with the closest values to the regular regression model and at the same time have the highest R^2. However, it should be noted that this method is used to derive equations that are very similar in form with Ou's and at the same time fit the data well. However, they may still not be the perfect fit to the data since other underlying factors may be significant that have not been considered in Ou's models. Besides, because of the complexity in Ou's forms of equations, further predictions of the powers used in the equations are not studied. This may hold back the appropriate analysis according to the data gathered from this experiment.

Table 24: Model 1 is the fitted equation based on Ou's form, Model 2 is the fitted equation based on the approximation technique.

Ou's equations	**Ou's Model**			**Approximation**	
	Coefficients	**Ou's R^2**	**Our R^2**	**Coefficients**	R^2
$HL = a + bL^*$	a=3.9 b=-0.05	0.76	0.85	a=3.5 b=-0.03	0.85
$WC = a + b(C^*)^c \cos(h - 50)$	a=-0.5 b=0.02 c=1.07	0.74	0.59	a=1.19 b=0.01 c=1.07	0.59
$HS = a + b(100 - L^*) + d(C^*)^e$	a=11.1 b=0.03 c=-11.4 d=0.02	0.73	0.61	a=-5.1459 b=0.024 c=5.33 d=0.02	0.66
$AP = a + b[(\Delta C^*)^2 + (\Delta L/1.5)^2]^{1/2}$	a=-1.1 b=0.03	0.75	0.52	a=0.0003 b=1.135	0.48

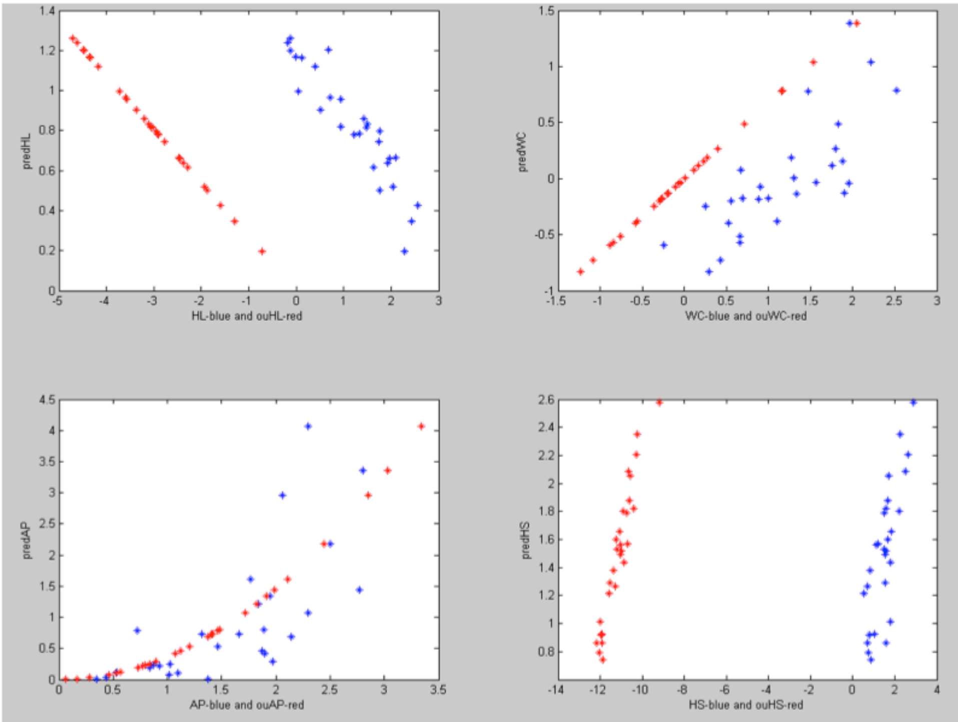


Figure 33: Scatter diagram for the four models with red points indicating predictions based on Ou's formulae and blue points indicating the predictions according to the data collected from the experiment.

5.4.4 Conclusion

The gathered data from the experiment show some correlation with values derived from Ou's models indicating that they may be predicting the data well. However the correlation values for some of the characteristics such as active-passive are not that high. Slight changes to the coefficients in Ou's equations, based on the gathered data may increase the correlation and initially improve the predictability of these models. Approximated models were generated according to this aim, resulting in new sets of formulae with the same form but with different coefficients. Additionally, throughout this process, the behaviour of the data is understood better.

In general it may be concluded that despite the differences between the methodologies of the two methods, the derived models may suggest that overall the variation across the data could be explained in a similar way. This further suggests that although different numbers of observers had taken part in the experiments (30 observers in Ou's experiment and 20 in this experiment) the results are very similar. Advising raising the number of participants may not be required. However, since crucial variables such as age and cultural differences did not show any significance it is required to run an experiment with high number of participants (with various age and cultures) in order to find the possible correlations (this clarifies the reason why a high number of participants is required for the global Global Online Colour Survey). In addition, as remarked earlier, one of the major differences between Ou's experiment and this experiment is the colour samples used. Therefore the similarity of the results may suggest that no matter whether the colour sample would be physical or displayed on monitors, they would almost be rated in the same way in terms of the bi-polar characteristics. The use of physical colour samples seems to indicate no advantage over the colour samples viewed on the LCD monitors. It also illustrates the fact that experimenter for each trial. Apparently, with the monitors being used widely across the globe and the ease of distributing such experiment is required to be present for each trial. It would appear, therefore, with the outcomes of this chapter can validate the Global Online Colour Survey.

6 Model fitting strategy

6.1 Introduction

Modelling the relationship between colour and bi-polar characteristics is the main subject of this thesis. There are various ways in which models could be useful. For example, they could be used as a basis for content-based image search on search engines such as Google (CBIR) or they could be used to develop tools that could assist designers in making colour decisions and/or teach design students about effective colour use. Models derived by Ou, for example, describe the bi-polar characteristics as a dependent to the colour values. In other words, the characteristics of a colour are predicted from the LCH values by the model. Recently an application has been created that seems to be based upon Ou's models and is available for the Android market. Figure 34 displays screenshots from this application. The designer would use the RGB or L*a*b* slider bars to select a specific colour as an input; the characteristics of that colour are then displayed on the second page[75].

Figure 34: Screen shots of the colour design assistant application on Android Market.

Although this application can help designers in understanding the characteristics for a specific colour, it may not be entirely useful for the colour selection process of design. The design process involves the identification of the effective factors that underlie the product's characteristics. Initially, the designer would have to find a colour that best suits these characteristics. For example, a designer would be looking for a colour that has certain

[75] This app also functions on the semiotics for the combination of colours. The third and fourth screen shots are relevant to this.

characteristics such as active, hard and warm. So a tool or app, which takes the desirable characteristics as the inputs and generates colours as outputs, would be more useful. Figure 35 illustrates the interface of a colour selector tool. In this tool, the user selects different values for the bi-polar characteristics, the tool generates several colours that are consistent with that vector of characteristics and the colours can then be applied to images of selected products to aid the designer in the visualisation of the colour application.

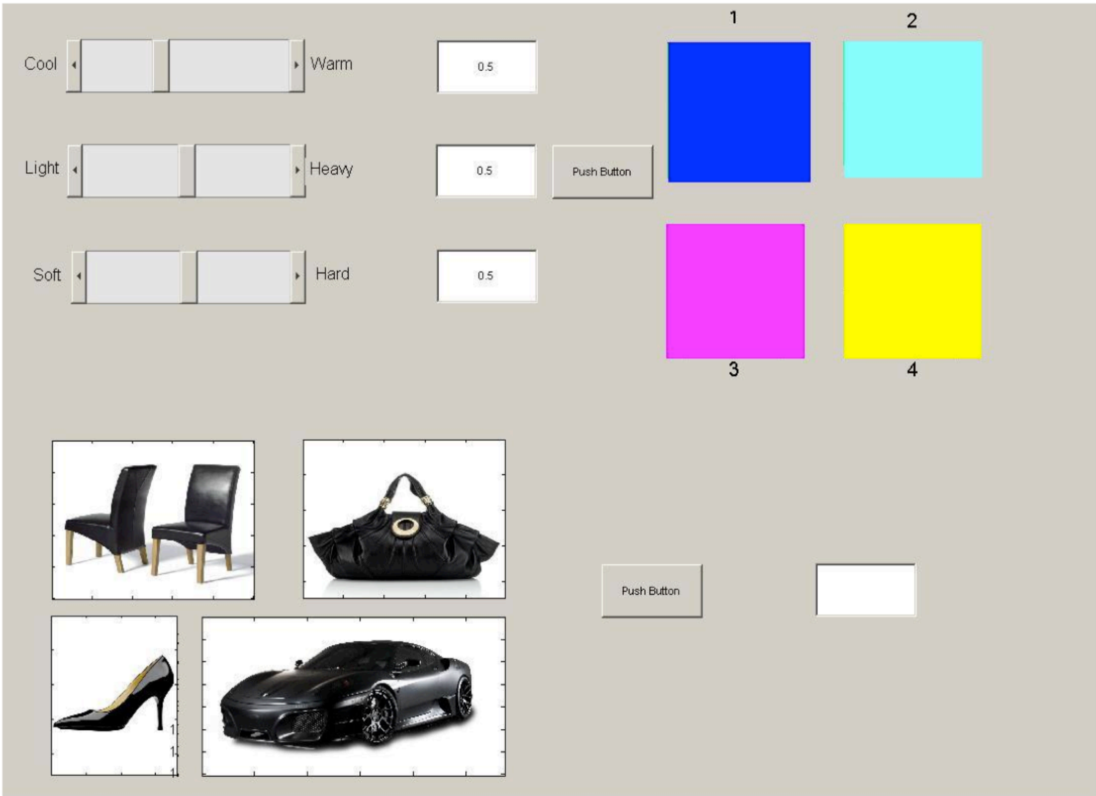


Figure 35: Interface of a colour selector tool with bi-polar characteristics as input and colours as output with the illustration of the selected colour on different products (leading to various colour perception upon application).

This colour tool would require the inverse relationships to those described by Ou's models. The problem is that it is not always trivial to derive an inverse relationship. In order to have models that are able to report the relationship between the input and output and also its inverse, simple equations are required that are reversible. In this chapter, the data gathered from the laboratory experiment are used to find invertible models describing the relationship between colours and bi-polar characteristics.

6.2 Methodology

In the majority of practical situations one may wish to specify the bi-polar characteristics and then predict a colour that would have those characteristics. This raises the question of why Ou's models should predict characteristics from colour rather than the other way around. The first reason is that there is a natural direction of causality; that is, the colour causes the meanings to be evoked – the meanings do not cause colours. Therefore colour is the natural independent variable. The second reason, which probably results from the natural direction of causality, is that several different colours are likely to have the same set of bi-polar characteristics. The data that were collected (in the last chapter and in Ou's work) constitute examples of mappings between colour space and, for example, a one-dimensional heavy-light space. Is it possible to use these examples to derive a mathematical model for the mapping that predicts colour from the bi-polar characteristic? This direction of mapping is considered to be difficult because it is likely that the mapping between colour and bi-polar characteristic will be $many \rightarrow 1$. Simple linear models describe a one-to-one (1→ 1) relationship such as $Y = a + x$, in which each value for *x* is linked to a single value of *Y*. Any equation with a degree higher than one, such as $Y = a + x^2$ forms a many-to-one ($many \rightarrow 1$) mapping. Now, more than one value of *x* can result in the identical value for *Y*. In our case, this means that two or more distinct colours (points in the colour space) map to a single point in heavy-light space. Conversely, this means that the mapping from heavy-light space to colour space will be a $1 \rightarrow many$ mapping, which means that there may not be a unique solution. However, $many \rightarrow 1$ mappings are easily modelled. (As an example, consider $Y = x^2$; this is a $many \rightarrow 1$ mapping between *x* and *Y*. The inverse relationship $x = y^{0.5}$ is $1 \rightarrow many$). Since, two values of *x* correspond to each *Y*, the reverse relationship between *Y* and *x* cannot be uniquely identified. Initially, solving the inverse problem becomes impossible when *Y* is relative to more than one variable such as $Y = a + x_1 + x_2^3$. It is for this reason that models that attempt to predict colour from bi-polar characteristics are not considered in this research. Rather, models that predict bi-polar characteristics from colour are developed. This is consistent, of course, with the work of other researchers in the field such as Ou and Sato [1, 160]. A method of solving what is known as the inverse problem will therefore be required. This involves some form of iteration or direct inversion of the models. In addition, when identifying the relationships, clarification of the interaction between the bi-polar characteristics is required in order to identify the correlation between the bi-polar characteristics[76]. Considering this fact also defines the regulations for modelling the relationship between bi-polar characteristics and colours, which is formally investigated in this chapter.

Initially, two different ways of approaching the problem of invertability are considered. One is to deal with the existing models. In general, they cannot be reversed since they tend to be non-linear. Therefore, approximation techniques could be used in order to investigate models that can represent the inverse relationships between colours and bi-polar characteristics. The second approach is to generate linear models according to the data collected from the experiment described in Chapter 5. Initially these models are inverted to conclude the relationship between colour and bi-polar characteristic in the reverse order.

6.2.1 Linear models

The best forms of invertible models are simple linear models. However, linear models do not always fit the data very well and therefore there is a potential trade-off between having a model that is simple enough to be inverted and yet complex enough to adequately describe the relationship. Take the data displayed in Figure 36 as an example. The question is when a large amount of data is considered which contains noise and outliers, what curve or line would best fit the data? A high-degree curve which absolutely covers all of the noisiness of the data with the minimum value of error or a lower-degree curve or even a line which fits the data less well?

The answer can be dependent upon the subject but there are some benefits in each approach. Although fitting a higher-degree curve that perfectly fits all of the collected data may be very accurate, it may not be very predictive. Whilst on the other hand, if the data are to be fitted with the lowest possible degree of curve, the true relationship may not be adequately described. Simple linear models can be a good general fit to the data as well as being invertible, since they are stable, do not include noise, more likely to fit the similar data and easier to inverse.

---

[76] Note that the interaction of the bi-polar characteristics has not been considered in previous research.

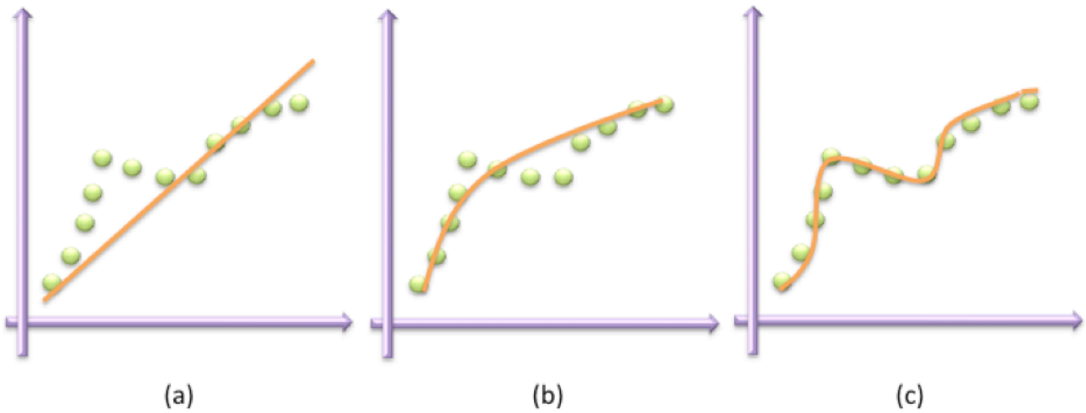


Figure 36: Fitting a curve to data; (a) a linear fit, which explains the overall relationship, (b) illustrates higher degree model (c) very accurate model that encounters the noise as well.

Bi-polar characteristics are typically explained by three variables; Ou's models define colours by LCH values. Hue as a variable has a different structure of variation in comparison to lightness and chroma. Since it is a polar coordinate that has a circular nature, Ou used the conversion $\cos(H - 10°)$. This means that lightness, chroma and hue do not vary in a similar way and this adds extra complexity to the structure of the models (resulting in them not being invertible). One-way of overcoming this problem is basing the models upon the variations of L*a*b* values rather than LCH, since all three components of L*a*b* vary in the Euclidean space and are almost uniform. In the next section different simple linear models based on LCH and L*a*b* variables are fitted and compared together. The main aim is to find the simplest models that are invertible.

The method of backward regression is applied to the LCH and L*a*b* values in order to find the relationship for all the ten bi-polar characteristics. The results are explained in the next section. In order to visually identify any possible relationship between the LCH values and bi-polar characteristics it is useful to look through the scatter diagrams. For example, heavy-light scatter diagrams are displayed in Figure 37. It is evident that lightness is highly correlated with heavy-light. It is expected that there would be a negative coefficient in the model that indicates this relationship.

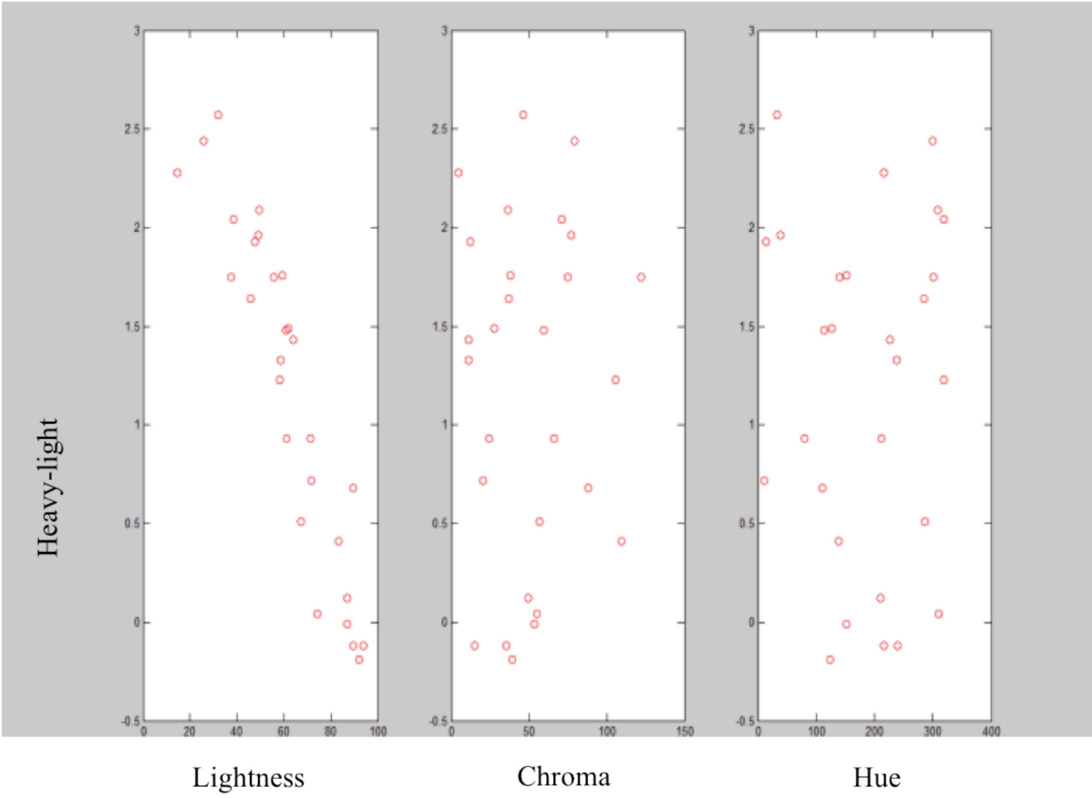


Figure 37: Scatter diagram for heavy-light

The method of generalised linear models is applied to the ten bi-polar characteristics. However, since each have a normal distribution and are not correlated to each other, it is best to apply the General Linear Models method. Backward regression technique has been carried out using SPSS and the coefficients for each model are displayed in Table 25. Table 26 contains the values for R^2, adjusted R^2. The model with higher value of R^2 (or adjusted R^2 for multivariate models) and with significance level less than 0.05 is chosen as the perfect fit to the heavy-light data.

Table 25: Table of coefficients for the regression models based on lightness, chroma and hue.

Model		Coefficients		Sig.
		B	Std. Error	
1	(Constant)	3.787	.245	.000
	Lightness	-.038	.003	.000
	Chroma	.001	.002	.511
	Hue	-.002	.001	.017

Table 26: Heavy-light model validation and the goodness of fit test.

Model	R^2	Adjusted R^2	Std. Error	Change Statistics			
				R^2	df1	df2	Sig.
1	.883	.869	.30908	.883	3	24	.000

The following equation is derived for heavy-light indicating a negative coefficient of 0.038 for lightness, 0.001 for chroma and -0.002 for hue. The adjusted R^2 value of 0.869 indicates that the model covers almost 87 percent of the variability across the data.

Equation 67

$$HL = 3.787 - 0.038L + 0.001C - 0.002h$$

The same procedure has been carried out for the rest of the bi-polar characteristics and the models derived are displayed in Table 27. However, as mentioned earlier; the major criticism against modelling the relationship between the bi-polar characteristics and the LCH values in a linear way is the circular variation of hue. In case of using cosine conversions, the discontinuity between 0 and 360 degrees is addressed. Since, Ou has used the conversion of cos $(H - 50)$ we will apply the same here. The second column in Table 27 illustrates the models derived based upon this conversion. The significance of the derived models has been evaluated using the F ratio test. Value less than 0.05, indicate significant relationship between the variables. The adjusted R^2 is also used to report the amount of variation covered by the model. Therefore, R^2 value close to one reports better performance of the model.

Table 27: Table of linear regression models derived for the ten bi-polar characteristics with adjusted R^2; using the backward regression technique in SPSS (multiply values by 10^{-3}).

	Linear Regression Model 1 coefficients				R^2	**Linear Regression Model 2 coefficients**				R^2
	Cons	**L**	**C**	**H**		**Cons**	**L**	**C**	$cos(H-50)$	
HL	3787	-370	1	-2	870	3541	-37	0	244	870
WC	2672	-140	0	-10	340	2126	-9	-5	910	780
MC	0	8	7	0	420	0	7	8	0	380
CD	0	150	10	0	470	0	14	13	0	550
AP	1274	0	11	-1	450	0	9	16	0	560
HS	2965	-20	6	-1	710	2695	-23	5	0	660
TR	2431	-10	10	-2	590	2102	-14	8	0	520
FS	2431	-10	10	-2	590	0	14	15	0	690
MF	0	0	0	0	100	2389	-14	0	-458	300
LD	1631	0	0	2	200	0	0	0	0	90

According to the adjusted R^2 values reported for Model 1, heavy-light and hard-soft have the best performance amongst all the ten bi-polar characteristics. Also, the masculine-feminine model is the only one that has not been identified as significant; therefore this model's coefficients and constant intercept are reported as zero. According to the F test, the rest of the models are found to be significant. However, the reported R^2 values are not very high in some cases. Also, in most cases the significant coefficient values are not reported any higher than 0.02 and some are as small as 0.001. This means that although there are significant relationships between the LCH values and each of the bi-polar characteristics, the data are not distributed linearly around a line with a considerable slope. Therefore the models would have small coefficients that are almost zero. Perhaps surprisingly there is not a great deal of difference between models that use hue and those that use cos $(H - 50)$ with the exception of the warm-cool model. This may be because hue does not actually seem to be a major factor in the majority of the models (warm-cool and masculine-feminine being exceptions to this).

It is also useful to compare Ou's models with the linear regression models derived in this section. Table 28 illustrates this comparison using the R^2 values of both models. It is very interesting to note that the performance of heavy-light and warm-cool are higher with the linear models rather than Ou's form. On the other hand, models for active-passive and hard-soft perform less well than Ou's. This may suggest that for these two bi-polar characteristics, linear regression may not fit the data well enough in order to explain the variation across the data. But overall, it should be noted that linear models have shown more significant results across the 9 bi-polar characteristics whilst Ou only found 4 significant models that explained the variation across the data.

Table 28: Comparison of the model based on L, C and $cos(h-o)$ with Ou's models.

Dependent variables	R^2 of Ou's models	Adjusted R^2 L, C and $cos(h-50)$
Heavy-Light	0.76	0.869
Warm-Cool	0.74	0.775
Active-Passive	0.75	0.562
Hard-Soft	0.73	0.662

However, the downside of this approach is the non-linear conversion of the hue that will prevent the models from being invertible. Therefore, application of L*a*b* values may be better since they vary uniformly in the Euclidean space. Following table contains the coefficients of the derived models for the 10 bi-polar characteristics with the relevant R^2 values. It appears that the models derived for heavy-light and warm-cool still have the highest performance. No significant models have been identified for modern-classic, active-passive and like-dislike. However, the rest of the significant models have considerable coefficients. Also, overall the variables lightness and b^* seem to have more significant impact across the ten bi-polar characteristics in comparison with a^*, since most of the derived coefficients of the ten models are zero or close to zero.

Table 29: Linear regression model with independent variables L, a and b.

Bi-polar characteristic	Linear regression model coefficients				Adjusted R^2
	Constant	L*	a*	b*	
Heavy-Light	3.631	-0.04	0.001	0.005	0.887
Warm-Cool	1.697	-0.008	0.014	0.014	0.716
Modern-Classic	0	0	0	0	0.114
Clean-Dirty	0	0.020	0	0	0. 32
Active-Passive	0	0	0	0	0
Hard-Soft	3.134	-0.026	0	0.004	0.679
Tense-Relaxed	2.522	-0.014	0	0.007	0.431
Fresh-Stale	0	0.019	0	0	0.184
Masculine-Feminine	2.995	-0.24	-0.016	-0.006	0.538
Like-Dislike	0	0	0	0	0.147

The advantages of using the models is that all three variables would linearly vary in the Euclidean space. However, it seems that the performances of these models are less informative than the models based on L, C and $\cos(H-50)$ as illustrated in Table 30. On average, the 10 models based on the L*a*b* variables are 28 percent descriptive, whilst the models based on LCH variables are 54 percent (see Table 31).

*Table 30: Comparison of the model based on L*a*b* with Ou's models.*

	Adjusted R^2	
Regression model	L, a and b	L, C and $\cos(H-50°)$
Heavy-Light	0.887	0.869
Warm-Cool	0.716	0.775
Modern-Classic	0.114	0.379
Clean-Dirty	0. 32	0.551
Active-Passive	0	0.562
Hard-Soft	0.679	0.662
Tense-Relaxed	0.431	0.515
Fresh-Stale	0.184	0.686
Masculine-Feminine	0.538	0.296
Like-Dislike	0.147	0.094

*Table 31:Descriptive statistics of the adjusted R^2 values for models based on L*a*b* and L, C and $cos(H-50°)$.*

	N	Minimum	Maximum	Mean	Std. Deviation
L*A*B*	10	.00	.72	.2809	.28447
L, C and $\cos(H-50°)$	10	.09	.87	.5389	.23220

Amongst the significant models derived, the equations in Table 32 have the highest adjusted R^2 values.

Table 32: Significant linear models.

Models
$HL = 3.631 - 0.04L + 0.001a + 0.005b$
$WC = 1.697 - 0.008L + 0.014a + 0.014b$
$HS = 3.134 - 0.026L + 0.004b$
$MF = 2.995 - 0.24L - 0.016a - 0.006b$

It is interesting to point out that Ou also found four significant models for heavy-light, warm-cool and hard-soft and active-passive. In general it can be concluded that only one or two of the models are accurate. Therefore using simple models to allow invertibility may result in too much loss of accuracy.

6.2.2 Inverse models

Up to now, at least three different types of models have been discussed and compared. In this section, each model is investigated in terms of its flexibility towards being reversed. Without any doubt, inverting a linear or a non-linear model has been the subject of much research up to now. The so-called inverse problem is a pure mathematical problem which is approached using complex algorithms and programmes. As mentioned earlier, mathematically inverting models based on lightness, chroma and $cos(hue - 50)$ and also Ou's models would be difficult. In general, inverting multiple variable regression models can be done using the inverse of the coefficient matrices. However, this matrix should be a square matrix in order to be invertible. On the other hand, since more than one independent variable may be involved in the model, it would have a $many \to 1$ mapping nature, resulting in a $1 \to many$ mapping for the reverse. Some of these models would not have a square matrix of coefficients; therefore approximation techniques should be applied. Commands such as *pinv* or backslash in MATLAB, compute a pseudoinverse of the matrix [11]. However, this command may only result in numerical values for one of the possible combinations of the multiple variables. There could be all sorts of other numerical combinations of variables, which give the same results. However, this technique is further explored with sets of numerical values that are used as inputs for the bi-polar characteristics resulting in sets of values as outputs for the independent variables. For example; consider the model for heavy-light (with the equation $HL = 3.631 - 0.04L + 0.001a + 0.005b$); a value for heavy-light is derived given values for lightness, a and b by solving the following matrices

Equation 68

$$HL = [3.631 \quad -0.04 \quad 0.001 \quad 0.005]\begin{bmatrix}1\\ L\\ a\\ b\end{bmatrix}$$

which alternatively can be displayed as

Equation 69

$$HL = AX$$

in order to find the values for matrix *X* which by the multiplication of *A* would result a specific value of *HL*; the following equation can be solved in MATLAB:

Equation 70

$$HL/A = X$$

Thus, in order to find the reverse relationship, a value for heavy-light is input and by using the backslash command in MATLAB, numerical values for lightness, a and b are derived. However, this technique may not be entirely reliable as it only provides one single response. The backslash technique uses an optimisation strategy, which functions until it finds a suitable response leaving many other possible results unidentified. Therefore, it would give a unique response even if repeated. For single variable linear models such as clean-dirty; it could be repeated with a large range of values, resulting in numerical sets for the lightness. It would be interesting to find out if these sets of values can help in finding the relationship between lightness (as dependent variable) and clean-dirty. However the major limitation to this approach is that it can only be applied properly to linear models that are not multivariate. Since transforming them into linear combinations of new variables, can affect the general nature of the variables. For example, the equation for warm-cool derived by Ou can be transformed into a linear model using the logarithmic transformation as follows:

Equation 71

$$\log\left(\frac{WC + 0.5}{0.02}\right) = 1.07\log(C^*) + \log\left(cos(H - 50)\right)$$

Using the method explained, a range of $C^*$and cos $(H - 50)$ can be derived for a set of WC values ranging between -2 and 2. But first of all these values are not unique and other numerical combinations of $C^*$and cos $(H - 50)$ can results in the same values of WC. Also, the logarithmic transformation would affect the natural variation of the data, which can be misleading.

Therefore, in order to find values for L*, a* and b* which correspond to a specific bi-polar characteristic scale, it is required to apply a numerical method in which trial and error would eventually make way to the correct response. An adaptive gradient descent method has been used and is explained in the next section.

6.2.2.1 *Inverting the existing models using approximation techniques*

A stepwise method is applied to the L*a*b* values in order to find the closest corresponding characteristic values using a process of gradient descent in error space [184]. Below is a step-by-step procedure of the algorithm:

1- L*, a* and b* values are randomly selected.

2- Using the forward model, the corresponding bi-polar characteristic value is derived.

3- The difference between the calculated characteristic value and the target value is derived.

4- A very small unit (0.05) is added to the one of the L*, a* or b* value, resulting in a different value for the bi-polar characteristic. The difference between this value and the previous error is calculated. Here, this difference will be named as additive-difference.

5- The same procedure is carried out, but this time by subtracting a small unit from the L*a*b* values.

6- A comparison between the additive and subtractive-difference is made, setting the smallest value to the best difference.

7- If either the increase or decrease in the value results in a lower error than the previously best error, then the increase or decrease is accepted otherwise it is rejected.

8- If the change (increase or decrease) accepted the step size for that value is increased otherwise it is decreased.

9- Steps 4 to 8 are repeated many times (each time on either L*, a* or b*) until one of several stopping criteria (such as the error reaching a specified level) are reached.

Once the smallest best-difference is found in step 9, the corresponding L*a*b* values would indicate the attributes of a colour with the closest characteristic value. However, at this point some issues arise in terms of colour gamut. The L*a*b* values generated by this algorithm are not constrained; even if the solutions are reasonable L*a*b* values they may be outside of the gamut of sRGB (which is smaller than CIELAB of course) and so when converted to sRGB space the RGB values may be invalid. Therefore, gamut mapping is required at this stage.

6.2.2.1.1 Gamut mapping strategy

Gamut mapping of the L*a*b* values can be extremely complicated and has been the subject of much research [185]. The approach taken in this study is to perform a simple gamut mapping in sRGB space; that is, to ensure that invalid RGB values are appropriately mapped to valid RGB values in the sRGB space. The sRGB colour space is in form of a three-dimensional cube with the values ranged between 0 and 1 (see Figure 38).

Typical sRGB conversion functions, replace the values less than 0 with 0 and values more than 1 with 1. This however, may cause some dramatic shifts, for example; say the converted RGB values of a colour from L*a*b* space is [0.3 3.7 0.9]. If the normal sRGB mapping is applied it will result in the RGB values of [0.3 1 0.9] which generates the colour displayed in Figure 39. Whilst if the RGB values are all scaled down equally by the division of all of the RGB components in the value which is higher than 1; then the result would be of RGB values equal to [0.0811 1.0000 0.2432], displayed in Figure 39. By simple visual comparison of the two images, it is very clear that the two methods result in complete different hues. Therefore, it is very important to note the correct gamut mapping technique[77].

[77] Note that the figure illustrates the effect of two different gamut mapping techniques that are broadly referred to as gamut clipping and gamut compression.

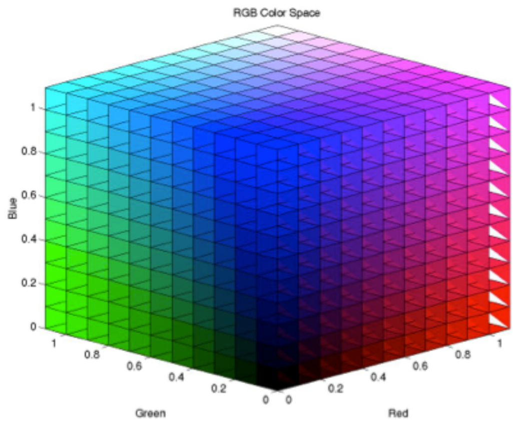


Figure 38: RGB three-dimensional colour space [186].

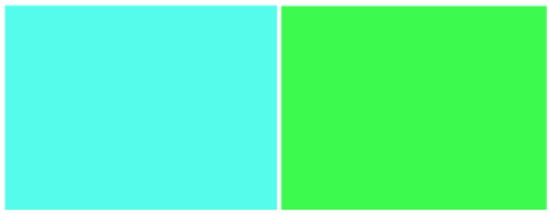


Figure 39: Colour with the RGB values of [0.3 3.7 0.9] mapped to [0.3 1 0.9] on the left and [0.0811 1.0000 0.2432] on the right.

The logic behind this scaling is in the structure of the colour cube. The RGB values of [0.3 3.7 0.9] are located outside the colour cube. In order to scale down all the R, G and B components together so that the colour falls in to the RGB cube, it is required to pull back the colour located outside the cube towards [0 0 0] on the imaginary line which connects them together. For example; the RGB value of [0.3 3.7 0.9] can be pulled back towards [0 0 0], on the line that connects them only if its individual components are divided by the element which is valued over 1. By this the main proportions of the individual R, G and B values would remain the same and therefore can form a reliable way of mapping. For cases where mapping is required for the negative RGB values; subtraction of the whole vector from 1 is applicable.

6.2.2.1.2 Results

In order to illustrate the results of the optimisation technique, a tool was designed in MATLAB displayed in Figure 40. In the example shown, various colours are generated for a heavy-light value of 20 and the sRGB representations of these colours are shown. Note that there are many different colours that all have a heavy-light value of 20 (which one is found depends upon the random starting point in the optimisation algorithm. (Note that the fact the algorithm is capable of finding multiple solutions could have positive benefits for a design application since the set of solutions could be used to generate a palette of colours that have a specific characteristic.)

Multiple solutions are inevitable; since colours vary in a three-dimensional space and tens or hundreds of L*a*b* combinations may result in the exact same value of a bi-polar characteristic. A consequence of multiple solutions is that if a user (see Figure 35) made small changes to a slider bar defining a characteristic (e.g. heavy-light) and the colour displayed in real time, then the colour could be seen to change quickly and markedly despite only small changes in the slider bar. This is because for each value of the characteristic the optimisation algorithm could find a very different solution. Phenomenally, the user may experience the colour to be flickering or flipping and therefore later this will be referred to as the flipping phenomenon.

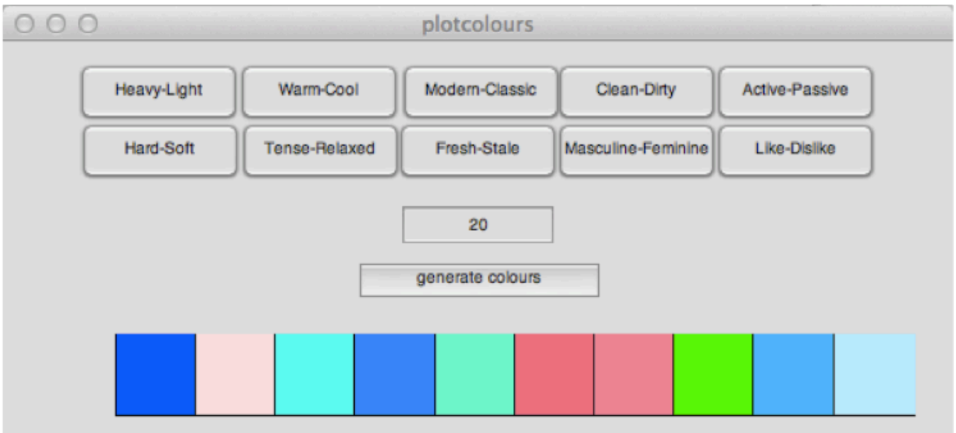


Figure 40: Tool designed to illustrate ten different results of uphill gardening method on each of the bi-polar characteristics in MATLAB. In this figure the scale value of 20 is chosen for heavy-light.

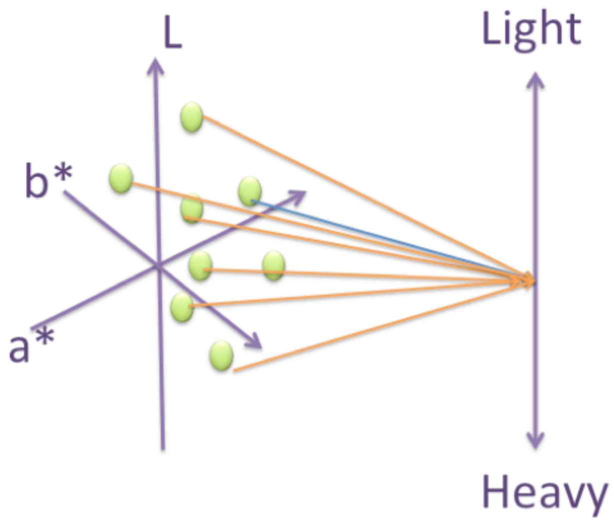


*Figure 41: Illustration of the many to one mapping from the three-dimensional L*a*b* space to the one-dimensional heavy-light space.*

Some of the linear models derived in Table 29 show significant relationship with only one or two attributes of colour such as the models derived for clean-dirty, fresh-stale or hard-soft and tense-relaxed. Despite the fact that inverting these models would be easier because of their simplicity; the inverse of these models may not be fully informative about colours. Since it would provide partial information about the colour; for example, inverting the clean-dirty or fresh-stale models would only provide information about lightness levels. Thus, again the simplicity of the former models takes the accuracy of the inverse models away. However, since lightness is the only effective attribute on cleanness or dirtiness of a colour, the variation of other attributes would not be effective. Therefore, these models and their inverse are still representative and could be used when other attributes of colour are already known, and only the value for lightness would be derived from the models as a significant factor on cleanness or dirtiness of that colour. Therefore, in order to find values for all three colour attributes, other models should be generated describing the inverse relationship based on L*, a* and b*. Methods for deriving these complimentary models are explained in the next section. Also another interface was designed in order to illustrate the progressive variation of colours across a scale from -2 to 2 (Figure 42). This represents the flow of colours from one value of a bi-polar characteristic to another. It is expected that a logical flow would be demonstrated across the nine squares illustrated in the figure. But unfortunately, this flow may be interrupted by unexpected colours because of the flipping effect, which mainly occurs because of the $many \rightarrow one$ nature of the models described earlier. Also, this tool would generate different range of colours each time it is re-generated.

Figure 42: Interface for the colour range with the heavy-light scale between -2 and 2.

These tools were generated based upon the inverse models of the L*a*b* linear models to investigate and illustrate the optimisation technique and its characteristics in a simple way. One of the advantages of the optimisation technique is that it is applicable for non-linear models as it is a numerical method and runs by trial and error. Thus having this method explained and illustrated for the linear models, it can be applied similarly to the non-linear models. However as mentioned earlier, there will not be a unique answer, which may be a result of the $many \rightarrow one$ mapping nature of the models and also because of the squared variables that exist in non-linear models. These variables can have two values, which lead to the same response, therefore cause flipping preventing a perfect flow in the sequence of the response. Also, note that the non-linear models are based on LCH variables; thus the structure of the variation would not be of the same nature adding more complexity to the method. Application of the approximation technique to the linear models would be adequate however, its application on Ou's models and other non-linear models based on LCH or $LC\cos(H-50)$ is carried with the very same procedure.

6.2.2.2 Numerical comparison of the approximation technique

In order to investigate this method deeper, 20 trials have been made to generate the heavy-light values equal to 1.5. As mentioned earlier, the derived L*a*b* values are not equal to each other. In order to form a comparison, additional 20 trials of approximation technique have been carried out which are illustrated in Table 33.

In order to compare the two sets of L*a*b* values, it is useful to calculate the ΔE_{2000} values for each of the two by two selections of the two sets. In other words, there are 20x20 possible comparisons, which can be made between the individual L*a*b* values of the two sets. The minimum value of the ΔE_{2000} illustrates the smallest difference between the colours generated by the approximation technique; in other words, pointing out to the most similar colours. By repeating the entire process 4 times, 5 different sets of ΔE_{2000} values are generated and compared in Table 34.

*Table 33: Two sets of generated L*a*b* values for heavy-light equal to 1.5, using the approximation technique.*

	L_1^*	a_1^*	b_1^*	L_2^*	a_2^*	b_2^*
1	57.53	15.46	30.93	43.79	-89.93	-57.89
2	60.48	32.45	51.12	53.52	0.10	1.97
3	52.03	5.84	-11.10	59.74	40.80	43.54
4	52.81	18.91	-7.51	48.90	68.98	-48.82
5	59.07	26.71	41.03	52.60	-97.17	14.03
6	55.86	0.06	20.64	29.17	-144.84	-163.86
7	58.22	18.33	35.93	58.01	-16.11	41.13
8	58.79	-61.98	56.53	19.53	-28.09	-264.38
9	38.59	-412.05	-35.06	54.05	-16.80	9.57
10	58.62	34.64	35.81	43.81	21.52	-79.99
11	39.31	-0.53	-111.62	25.19	27.87	-230.27
12	54.62	33.88	3.95	56.30	5.63	23.03
13	55.25	10.32	13.72	52.23	-82.26	8.05
14	53.95	13.43	2.72	58.96	11.35	43.21
15	52.19	-30.59	-2.56	52.32	52.60	-18.20
16	55.52	40.52	9.89	58.80	29.51	38.33
17	56.55	20.09	22.18	57.22	-1.81	31.91
18	54.70	26.75	6.05	56.19	52.97	12.77
19	46.18	-15.99	-53.59	55.93	7.31	19.78
20	48.39	-184.74	-2.17	56.58	-37.79	34.02

Table 34: Minimum and Maximum ΔE_{2000} values for each of the trials.

Trial	1	2	3	4	5
Min ΔE_{2000}	2.65	1.22	2.11	1.57	1.30
Max ΔE_{2000}	109.03	138.75	128.85	144.91	132.49

By a brief look into this table, it is clear that the variation across the generated L*a*b* values is relatively high. However the minimum ΔE_{2000} values are no higher than 3 units. This explains that the colours generated by the approximation technique can range from being very similar (only 3 units of ΔE_{2000} difference) to very different (over 100 units of ΔE_{2000} difference) from each other. In order to test the equality of the L*a*b* pairs a t-test is required. The null hypothesis $H_0: \Delta E_{2000} = 0$ will be rejected if the two-tailed significance level is less than 0.05 concluding that the two samples are not equal.

Table 35: One sample t-test carried out on the 5 sets of ΔE_{2000} values.

	Test Value = 0					
	t	df	Sig.	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
DeltaE1	36.97	399	0.0	39.31	37.22	41.40
DeltaE2	35.46	399	0.0	40.73	38.47	42.99
DeltaE3	36.35	399	0.0	43.11	40.78	45.45
DeltaE4	34.19	399	0.0	43.13	40.65	45.61
DeltaE5	37.05	399	0.0	44.92	42.54	47.30

According to Table 35, the null hypothesis is rejected for all 5 samples. This means that, despite the 20 reputations, the optimisation technique never generates the same values for the L*a*b* values. This directly points out to the $many \rightarrow 1$ mapping nature of the models that was mentioned earlier in this chapter. On the other hand, it shows that the optimisation technique has the ability to generate most of the possible L*a*b* values that correspond to the heavy-light value equal to 1.5. Which is considered as one of the advantages of this technique since it generously provides possible results.

6.2.2.3 Discussion

An adaptive gradient-descent optimisation technique has been adopted to numerically solve the inverse problem for sets of linear models. However, the solution may not be unique since flipping occurs. This may be because of the $many \rightarrow one$ nature of the models and also because of the squared variables that exist in non-linear models. The advantage of the non-unique response of the up-hill gardening method is that by repeating the technique continually a full pallet of colours can be proposed that has the desirable characteristics. Thus rather than just presenting one colour to the user, a range of colours would be proposed. The user would have more options to choose from a range of colours that harmoniously illustrate a certain value for a bi-polar characteristic. Thereby, it is beneficial to take full advantage of the non-unique response of the optimisation process. There is, however, a negative aspect of the flipping; that is, a user moving a slider bar that changes the value of a bi-polar characteristic may experience flickering of the corresponding colour on screen.

Also note that, since no significant models were identified for modern-classic, active-passive and like-dislike; it would not be possible to identify the inverse relationship. Also, since some of the models do not show a significant relationship with all of the L*, a* and b* variables it would not be possible to generate a full colour[78]. This specifically occurs for clean-dirty, hard-soft, tense-relaxed and fresh-stale. In this situation there would not be enough information about the colour attributes. By applying additional information from the correlated characteristics parallel to inverting the models, it may be possible to build up a full colour. Another difficulty arises for the characteristics that do not have an identified significant relationship with L*, a* and b* values. If a characteristic is highly correlated with another (which is modelled significantly with L*a*b* values), then there could be ways of identifying meaningful relationships for that bi-polar characteristic respectively.

In conclusion; investigation for additional models is required, which may provide additional information about models which do not have significant coefficients for L*, a* and b* by providing additional information about the structure of the variation across the colours and

[78] However, the tool can be designed such that the user would be able to input some of the attributes of the colour manually. So the colours would generate partially by the input of the user and the inverse relationships.

the bi-polar characteristics. These models could be proposed based on the correlations between the bi-polar characteristics. It is predicted that this source of information may aid the inverse problem. Therefore a deeper investigation of this subject is described in the next section.

6.3 Conclusion

The main aim of this chapter was to investigate how to address the inverse problem for models that predict bi-polar characteristics from colour attributes.

Firstly, linear models based on L*, a* and b* values were derived for each of the bi-polar characteristics. Although these models were less accurate than non-linear models based on LCH values, they were selected for the application since they could be inverted more easily. Note that only 4 of the bi-polar characteristics were modelled with relatively high R^2 value. In general it was concluded that only one or two of the models were accurate. Therefore using simple models to allow invertibility may result in too much loss of accuracy. This is a trade-off between accuracy and invertibility and unfortunately the terms of the trade-off are probably not acceptable.

Secondly, a numerical method (adaptive gradient descent optimisation) was implemented to solve the inverse problem. The advantage of this approach is that it can work with linear or nonlinear models. In this chapter we developed the optimisation method using linear models. It was necessary to develop an sRGB gamut mapping algorithm to ensure that solutions were physically reasonable.

In conclusion, this chapter has enlightened the knowledge about modelling the two-way relationship between bi-polar characteristics which will later on be applied to the colour selector tool. This tool would be generated according to the algorithms described in this chapter but based upon models derived from the Global Online Survey data collection in Chapter 7.

7 The Global Online Colour Survey

7.1 Introduction

Colour's relationship with semiotics is the main subject of this thesis. A few researchers have formulised these relationships in the form of so-called colour semiotic models. Most models of colour semiotics are based on data obtained from very small (and often very specific) groups of participants. This study conducts a colour rating experiment (where participants rate certain colours in terms of certain meanings) using a web-based survey. The aim is to obtain thousands (or tens of thousands) of responses so that robust and meaningful models of colour semiotics can be developed. There are three specific objectives for this extensive data collection:

1. To determine robust estimates of colour semiotics;
2. To develop models that can predict colours that would optimally convey certain meanings.
3. To establish the web-based paradigm as a valid procedure for colour experiments of this nature.

Thus a survey was required whose collected response would allow deep analysis upon the existing formulas and also opportunity to create new formulae for colour semiotics and their relationships. For this reason a very large amount of data was required so that: first, it would cover age and culture differences amongst different people, and secondly, it would lead to a higher accuracy level for derived formulae. It is very important to consider that due to the huge variety of perception amongst different personalities, full accuracy of formulae would only happen in the case that most individuals in the world would be questioned; this, of course, is not practical. However, this is the first large-scale global experiment to attempt to determine colour semiotics quantitatively. Previous work [169] has been based upon robust quantitative assessments of small groups (typically less than 30) of participants who are unlikely to be representative of the wider population or based on unreliable contemplation of an individual's experience or beliefs. In addition, the methods used in this study (see below) may be able to address the related research question of whether colour semiotics are global or culture specific. But even though the Global Online Colour Survey may be widely distributed containing a wide variety of data, it may still not involve all the influential factors; so for example, the colours of this survey do not contain all the possible colours that exist in the visual world (it is based on just 28 selected colours) and the survey does not begin to address factors such as shape and texture, etc. Still it is the first step towards mass data collection on colour semiotics and can be re-developed in future studies with consideration of more factors.

It may be though that that the rise in the number of respondents is not effective as it arouses other issues. These could be variability of the viewing condition, non-calibrated monitors of different users, variability of the environmental lighting conditions, variation of distance between monitor and observer, even instability of the respondent's attitude and mood at the time of response and many other more variables which can make the data seem unreliable. However, although controlled conditions of laboratory experiments shape a precise set of data, the analysis of the data can only be valid for that specific situation. For example, if a certain set of colours is chosen to be displayed in the L*a*b* for a certain age group of respondents selected from a certain location (which could correspond to where they live), then the data collected from this sort of experiment strictly is only applicable to that set of conditions. Thus, data collected from Chinese respondents in the study by Ou [169] was used to develop colour semiotic models which correspond to that specific group of respondents with a shared cultural background, approximate age and language. This raises the question of whether these formulae are valid for all the people around the globe and useful for successful marketing or advertising [120]. Indeed, it can be argued that the adding noise to an experiment can actually be beneficial [2]. Reips quotes a rule-of-thumb proposed by Martin that "the more highly controlled the experiment, the less generally applicable the results. If you want to generalise the results of your experiment, do not control all of the variables. Of course, this does not indicate that one should not try to control the noise in an experiment at all. However, it indicates that some noise in the experiment may not be as detrimental as might first be thought. Additionally, a small experiment was carried out for the purpose of understanding the magnitude of colour appearance variation which occurs because of the un-controlled nature of the Global Online Colour Survey. There is no doubt that the participants of the Global Online Colour Survey will view colours in all sorts of environments in which display monitors, surround lighting conditions, viewing distance and even size are different. Although, we are aware that this difference is not huge since the science of colour management has made a huge improvement in the field of display technology so far. Thus, with the current display devices, it is very rare that a single colour, say red, appears yellow or green on another device. The experiment was carried out by measuring the appearance of colours with a spectrophotometer located in various typical viewing conditions measured in random order. The surround luminance was also changed randomly in order to simulate possible viewing conditions (bright, dim, dark and average lighting). As expected no significant hue shift occurred across the 21 conditions, however lightness and chromatic variety was evident. The results suggested that the displayed colours do not appear very different in terms of hue. Overall, the colours seem to appear lighter than expected. This is not surprising since when measuring the colours on various LCD displays, the white point of the LCD's are different, and also veiling glare effects the colour appearance (specifically when the surround luminance is average or bright)[79].

The other problem is the sample size. In most of the experiments, the sample size is limited by time, the L*a*b* restrictions, calibration, and one-by-one observations. Often, it takes time for the monitor's colours to get calibrated to the standards, whilst the respondent also needs some time to get adapted to the experimental situations such as room darkness. Therefore, because of the time and effort that each respondent needs to invest, previous lab-based experiments in this field are necessarily limited in terms of numbers of participants. Another advantage of the study in this thesis is that a wider range of variation can be considered, rather than collecting data from a very tiny group (see Figure 43).

With small-scale colour experiments, each individual normally makes decisions about all the selected colours. There is the possibility that the length of the session required could be excessive and this could lead to problems such as the participant becoming bored or frustrated[80]. However, in a web-based experiment where thousands of participants will be involved, it will not be practical to have each participant rating each of the colours (it is likely that a participant will only stay on the webpage for a minute or two). Therefore, a new experimental paradigm is needed where the experiment is distributed over a wide population. In this study no single observer will take part in the whole experiment (each will

---

[79] The conclusions are based on the analysis of ΔE values, which are not mentioned here because of space limitations.

[80] Whilst in the web-based approach; each observer would be presented with only one colour out of the full range making the assessment pure and isolated from the rest.

undertake a small fraction of the experiment); nevertheless, the cumulative effect of doing this for large numbers of participants can be effective (see for example [2]).

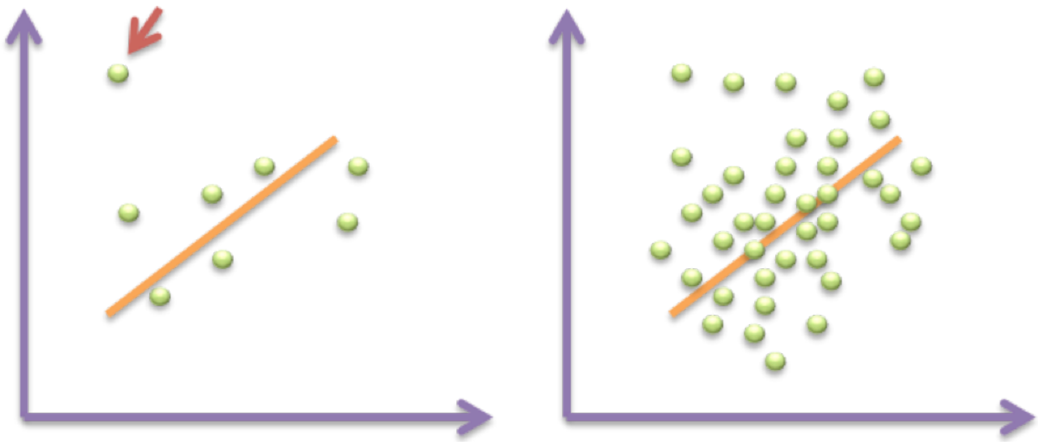


Figure 43: Right: small sample controlled experiment and the position of an outlier, left: web based experiment

One possible objection to this work is that about 5% of the population is so-called colour blind. However, there is no practical way to screen these people out in an on-line survey (since many people are not aware that they are colour-blind)[81]. But it should be remembered that even this 5% of the people buy products and are part of the population about whom we wish to deduce properties for; the idea of excluding colour-blind observers is much more critical if a small number of observers (typically less than 20) undertakes a Lab-based study since one or two colour-blind observers can make a big impact on the results.

Nevertheless, there remain some doubts about this data collection technique. Fortunately a critical study done by Beretta and Moroney (2011) [187] on colour naming indicates that large scale data, collected online, without any specific limitation correlate well with limited amount of collected data obtained in controlled condition such as laboratory. Their main intention was to perform a laboratory validation of the vast uncontrolled data collected from the web as they found it critical to compare the actual informal character of survey

[81] Introducing an on-screen test for colour blindness as part of the experiment would massively reduce the number of people willing to participate.

with the formal data collection process [188]. They targeted the Munroe and Ellis Lab-based experiment data which was collected online and did the same experiment in a controlled situation in the Lab. According to their results, data collected from both techniques were reasonably correlated to each other indicating that the two techniques work similarly. There are similar cases were data has been collected over a decade on a large scale through the web such as Moroney's colour naming survey [2]. Other recent studies suggest the excellent agreement for two different web-based experiments which was carried out upon the subject of colour naming [189]. Another advantage of this technique is that the results, or the data itself, can be accessed easily and analysed over even millions of participants [190]. It is true that low participation rates would lead to a waste of time for those who had taken part. On the other hand, those people would have wasted only about 2 minutes of their lives. In this study it is necessary to choose between two approaches – (i) get people to take part in a long on-line experiment (typically 30 minutes) for which we would have to reward them of course, and would even then struggle to get 50+ participants or (ii) get 1000s of people to take part, in which case we can't practically reward them but nor can we expect to take up very much of their time at all. This study has clearly adopted method (ii).

7.2 Methodology

Since participants are unlikely to spend more than a few minutes[82] on this on-line experiment, the experiment is partially distributed; that is, each participant takes part in only a fraction of the experiment. In addition, the experiment is presented in several languages (English, Chinese, Spanish, French, Italian, Urdu, Malay, Korean, Persian, and Arabic (see Figure 44)) since another issue with previous work is the reliance on English as the basis for the study even though the participants' English language may be limited.

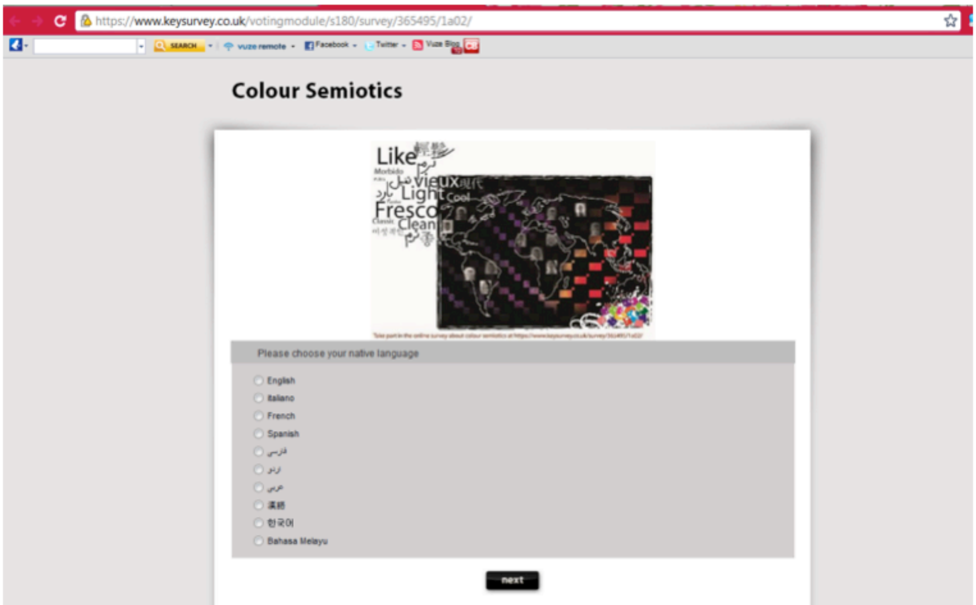


Figure 44: First page of the Global Online Colour Survey.

With reference to cultural differences, participants can respond according to their own opinion based on their own culture communicated within their own language. Obviously, for this study more languages could have been included but because of time and fund limitations ten were deemed sufficient. But at this point a new concern must be considered which is that the data will only be meaningful if the translations are definitely consistent

---

[82] Calibration was not found to be necessary as it extends the experimental time. Also this gathers the data from a real visual environment (e.g. online shopping) where all different kinds of displays are used.

with the English originally. There may not always be a one-to-one correlation for terms between different languages. However, this approach is certainly superior to what has traditionally been published in the past (whereby English language is used for all participants and some of the non-native-English speakers may not have a good understanding of the terms at all). However, to minimise the concern; translators were asked to undertake the work that satisfied all of the following three conditions: (a) native speakers of the non-English language; (b) good speakers of English; (c) some understanding of the field of research (colour and design).

The Global Online Colour Survey was programmed by the aid of a professional survey builder software offered by the Key Survey World App website [171]. In this survey, personal information (age, gender, nationality, native language)[83] along with participant's opinions on the relationships of certain colours with certain meanings are collected. The second and third page of the Global Online Survey are the very same as used in the lab-based experiment displayed in Figure 24 and Figure 25.

The same survey lay out and bi-polar characteristics have been used to be consistent with the lab-based experiment and most previous research. Each participant is presented with a colour patch (selected at random from 28 colours that are distributed throughout colour space) and asked to indicate (using a mouse click) whether, for example, he/she perceives this colour as warm or cool (or some point between these two extremes). The participants complete this information for all 10 bi-polar relationships but for one colour only (each participant visit brings up a random colour).

Since each participant only carried out about 1/30th of the experiment (unless they repeat the experiment in which they will be classed as a new participant anyway), the experiment requires about 30 times as many participants as an equivalent laboratory-based experiment, where each participant takes part in the whole experiment. The best of the published studies have about 50 participants which would lead to an estimation of about

---

[83] In this survey any information (such as name or address) that could identify any individual participant was not required. In addition, we are using software on a *https* secure server that means that the geographical locations of any participants are unavailable to us.

1500 participants. (We can suggest that 1500 participants in a web-based experiment would be more robust than 50 participants in a lab-based experiment.)

7.3 Promotion

The Global Online Colour Survey was only accessible for those who have Internet access. Recent statistics indicate that 32.7 % of the world's population have access to Internet [191]. According to Table 36, over 50 percent of the total population of America, Australia and Europe use the Internet. Therefore by collecting data online, not only it would be possible to reach for participants from all over the world that have access to Internet, but also to that part of population whom use the computers and view colours on computer displays.

Table 36: World Internet users and population stats updated on 31 March 2011 [191].

World regions	Population	Internet users	%Population
Africa	1,037,524,058	139,875,242	13.5 %
Asia	3,879,740,877	1,016,799,076	26.2 %
Europe	816,426,346	500,723,686	61.3 %
Middle East	216,258,843	77,020,995	35.6 %
North America	347,394,870	273,067,546	78.6 %
Latin America	597,283,165	235,819,740	39.5 %
Australia	35,426,995	23,927,457	67.5 %
World Total	6,930,055,154	2,267,233,742	32.7 %

However, having the survey online may not mean that all the Internet users would be attracted towards it. For this survey, a series of media campaigns have been used to publicise the experiment and drive traffic to it, so that a considerable number of participants would participate over a 12-month period. Certain ways of promoting the Global Online Survey were used such as providing a URL address and displaying it on public and social websites, blogs and forums [192-194]. Another way is to invite people to take part by e-mailing them an invitation to the survey, containing the survey's URL. The advantages of using Email are that it is cost effective, quick and the invitation distribution is well targeted and controlled [195]. For the purpose of this research, access to many different E-mailing lists of different groups and societies were provided. In addition, an online animated video-presentation was designed which had an inviting role as well as being informative [196]. Also a blog was created especially for promoting the survey online, by having a link to the survey, which soon became a subject of interest of many people. A regular update of the total responds to the colour survey was reported on the blog. Also, it gave a good opportunity for different people around the world to add their comments and even argue about the advantages and disadvantages of this approach. This lead to further discussions on other colour related blogs and websites, which grabbed attention of various people all over the world. As a result a few online interviews were published regarding to this subject [197]. The survey was also publicised in the website of the University of Leeds - School of Design, encouraging students for participation. Alongside all of the mentioned promotions, effort has been made to make the survey available to search engines. By allowing each search engine to index the survey so that it will not only appear in relevant search results but also appear as the first few results on the relevant page. This was achieved by selecting appropriate key words for the survey and blog. Also, online press releases were taken as a source for informing social readers about this survey which lead to the high ranking of general search engine enquiries. Google search engine as an example, lists the blog, survey itself and the interviews as the first few search results for the word "colour semiotics". Social media sites such as Digg, Twitter LinkedIn and Facebook were also used for raising awareness about the survey. Use of these sites was very effective such that within the first week of the launch the survey had over 100 responses. Also an animated video clip was made in order to promote the survey on the YouTube and was also introduced on forums.

The survey was live for a 12-month period resulting in the collection of 2,273 responses from around the world[84]. In the next section descriptive tables of the collected data are displayed along with subsequent analysis.

7.4 Results

Overall 2,273 data were collected online with over 58 different nationalities taking part. However, this figure does not include the incomplete responses. Since each of the 28 colours was proposed in a random order, the number of responses may not be equal for all the 28 cases. With a participation of more than 2000, on average each colour has been

[84] Additionally around 450 people left the survey incomplete and their data is discarded.

semiotically evaluated 80 times. Compared to the regular lab-based experiments which typically run with 30 participants, it may be considered as more accurate since each participant would focus on only one colour rather than a whole set of colours (and therefore would not become bored or tired). Also, because each participant is different from the other there is more diversity amongst the collected response.

Table 37: Total number of responses for the 28 colours with 2273 participants.

	Colour	Number of response
1		74
2		76
3		99
4		82
5		81
6		83
7		94
8		85
9		87
10		80
11		84
12		80
13		76
14		85
15		69
16		85
17		70
18		68
19		98
20		88
21		76
22		70
23		79
24		94
25		59
26		79
27		78
28		67

7.4.1 Z scores

Firstly, it would be useful to analyse the 2,273 data all together by taking a similar approach as the lab-based experiment analysis. Therefore the Z scores for each of the ten bi-polar characteristics are derived from the frequency matrix and displayed in Table 38.

In order to find out if the Z scores derived from the Global Online Survey are similar to the ones derived from the lab-based experiment a one-sample t-test was carried out on the difference of each of the bi-polar characteristics.

By testing the significance levels of the one sample t-test with the null hypothesis of $H_0: Mean\ difference = 0$, and also the 95% confidence interval of the differences for the ten bi-polar characteristic's Z scores it is concluded that the null hypothesis is accepted for active-passive, hard-soft, fresh-stale and masculine-feminine Z scores. Heavy-light, warm-cool, modern-classical, clean-dirty, tense-relaxed and like-dislike do not have significantly equal Z scores. However, if the reported Z scores are different it may be misleading since the variance of Z scores can depend on many different variables. So it is best to examine the correlation between the two reported Z scores for each of the ten bi-polar characteristics. Graphical presentation of this comparison is displayed in Figure 45 in which it is very clear that there is a strong positive correlation for each of the ten bi-polar characteristics. Thus, it can be concluded that both web-based and lab-based experiments produce Z scores that vary in a similar way and that both experiments are in agreement with each other.

Table 38: Z scores derived from the Global Online Survey data.

	Colour	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
1		1.21	1.76	1.89	2.21	2.12	1.49	1.55	2.04	1.25	1.49
2		2.56	2.28	1.22	1.31	1.41	1.84	1.72	0.89	1.69	1.44
3		1.85	0.70	1.45	2.62	1.79	1.96	1.55	1.86	2.05	1.39
4		2.25	1.39	1.19	1.93	1.54	1.48	1.13	1.58	1.88	1.44
5		1.02	0.65	1.69	2.67	1.40	1.10	0.78	1.97	1.89	1.72
6		2.04	1.33	1.05	1.09	1.04	1.47	1.40	0.63	1.89	0.80
7		1.08	1.34	1.92	1.99	1.80	1.15	1.45	1.73	0.52	1.21
8		0.49	0.45	1.94	2.44	1.52	0.90	1.04	2.00	1.31	1.53
9		2.37	1.75	1.44	1.71	1.56	1.61	1.42	1.31	1.06	1.60
10		1.78	0.98	1.75	2.04	1.38	1.42	0.99	1.46	1.75	1.81
11		0.63	0.75	1.92	2.28	2.00	1.24	1.33	2.06	1.58	1.46
12		1.43	1.41	2.00	2.36	2.03	1.60	1.69	1.90	0.36	1.21
13		2.77	2.29	0.97	1.24	1.20	1.81	1.49	0.75	1.80	1.28
14		2.38	1.56	0.94	0.88	0.66	1.38	1.16	0.61	1.99	0.99
15		0.80	0.36	1.78	2.76	2.01	1.39	1.28	2.20	1.49	1.68
16		0.37	0.52	1.29	2.72	1.26	1.29	0.99	2.16	1.26	1.70
17		2.59	0.95	1.27	2.29	1.40	2.14	1.29	1.61	2.39	1.73
18		1.14	1.01	2.12	2.95	2.59	1.81	1.70	2.77	1.41	1.51
19		1.95	2.21	1.73	2.02	2.73	2.00	2.58	1.71	1.56	1.55
20		2.12	1.18	0.96	1.75	1.48	1.57	1.04	1.42	1.91	1.20
21		2.20	1.00	1.43	1.45	0.73	1.80	1.33	0.54	2.41	1.20
22		2.29	0.85	1.41	1.94	1.15	1.92	1.02	1.45	2.08	1.81
23		2.31	1.72	1.45	1.84	1.70	1.52	1.37	1.46	1.03	1.83
24		0.70	1.66	1.49	2.43	1.59	0.95	0.94	1.76	1.03	1.44
25		3.11	1.17	1.44	1.80	1.29	2.90	1.75	1.32	2.33	1.92
26		1.27	2.12	1.65	2.00	1.79	0.86	1.18	1.89	0.81	1.29
27		1.85	0.85	1.37	1.75	0.82	1.25	0.93	1.41	1.82	1.54
28		1.74	2.33	1.72	1.99	2.13	1.58	1.74	1.70	1.34	1.43

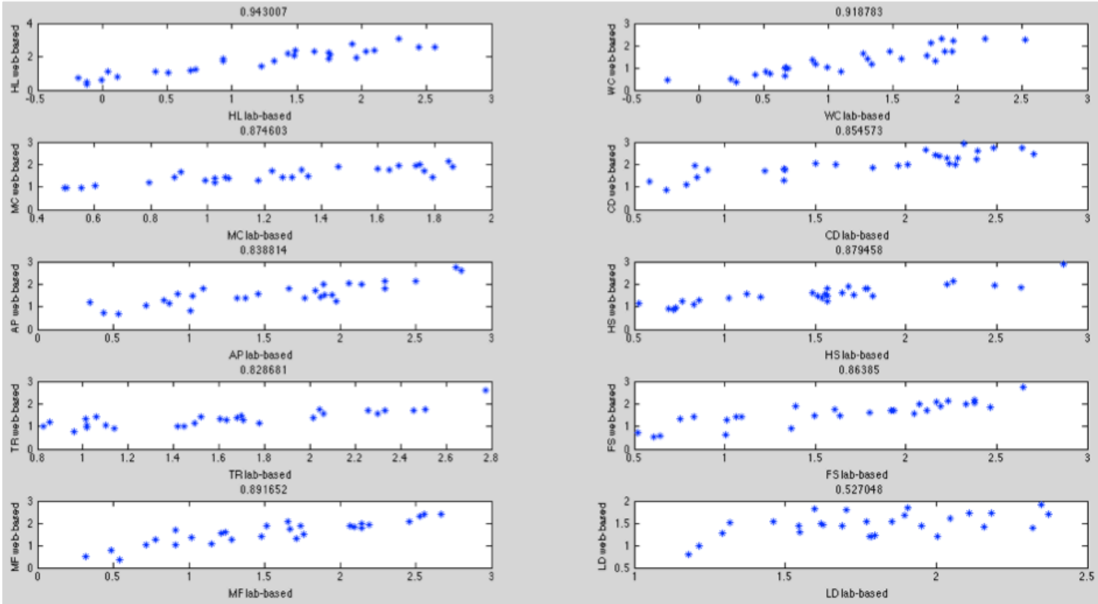


Figure 45: Scatter diagrams of the Z scores derived from lab-based and web-based experiments for the ten bi-polar characteristics.

But before any further analysis, it is important to check the participant agreement using the Kappa test in order to spot if respondents have been answering randomly for the 28 colours. Table 39 contains the Kappa test values suggesting that the only bi-polar characteristic which shows significant difference of rating amongst the 28 colour range is clean-dirty and active-passive. The agreement of the ratings amongst the 28 colours is significant amongst the other bi-polar characteristics. The results of the Kappa test are identical to the lab-based experiment suggesting that despite that the data collection has not been closely monitored or controlled, still the input has not been random. The observer accuracy values reported in Table 40 are not very high in value. This can be a result of having thousands of people involved in the experiment. In general it should be noted that since the observer accuracy formulae is not independent of the total number of participants and colours involved in the experiments, it may not be strong enough to be used for comparison. Specially, when it comes to comparison, this metric has a weak performance.

Table 39: Table of the Kappa test values for the Global Online Survey bi-polar characteristics.

	Kappa value	Var(Kappa)	Z
Heavy-light	0.1812	0.0	0.7654
Warm-cool	0.1477	0.0	3.2669
Modern-classic	0.0539	0.0	0.1981
Clean-dirty	0.1272	0.0	1.5211
Active-passive	0.0980	0.0	0.0701
Hard-soft	0.0615	0.0	0.4555
Tense-relaxed	0.0489	0.0	1.3182
Fresh-stale	0.1058	0.0	4.4110
Masculine-feminine	0.1174	0.0	0.1485
Like-dislike	0.0234	0.0	0.0803

Table 40: Table of observer accuracy values for the 10 bi-polar characteristics.

	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
Observer Accuracy	0.48	0.37	0.18	0.35	0.23	0.17	0.11	0.28	0.47	0.13

7.4.2 Correlation Matrix

A visual presentation of the correlation matrix between the bi-polar characteristics is displayed in Figure 46. Considering the absolute correlation values higher than 0.6 it can be concluded that fresh-stale is highly correlated with modern-classical clean-dirty and active-passive. Also active–passive is highly correlated with modern–classic. Clean- dirty is negatively correlated with heavy–light while positively correlated with modern-classic. So far, reported correlations are the same for the data collected in the lab-based experiment. However, additionally heavy-light is positively correlated with hard-soft and negatively correlated with fresh-stale and modern-classic.

The one-sample t-test comparison could be used in order to find significant similarity between the correlation matrix derived from the Global Online Survey and the lab-based experiment. According to the high significance values reported, it is evident that the two correlation matrices are significantly similar to each other. Therefore, it can be concluded that no matter how many participants take part in such experiment, the correlations between the bi-polar characteristics will be the same. This also proves that the Global Online Survey works as well as the lab-based experiment, which is carried out under controlled conditions. This fact can draw a line on the idea of possible bias in the Global data caused by the uncontrolled conditions. Bearing in mind that the Global Online Survey will provide more information regarding to cultural and age differences of the participants, it is preferred over the lab-based experiments. Overall, since the correlation values are mostly higher than 0.3, there is a likelihood that a factor structure underlies the variables and therefore further investigation is carried out through factor analysis.

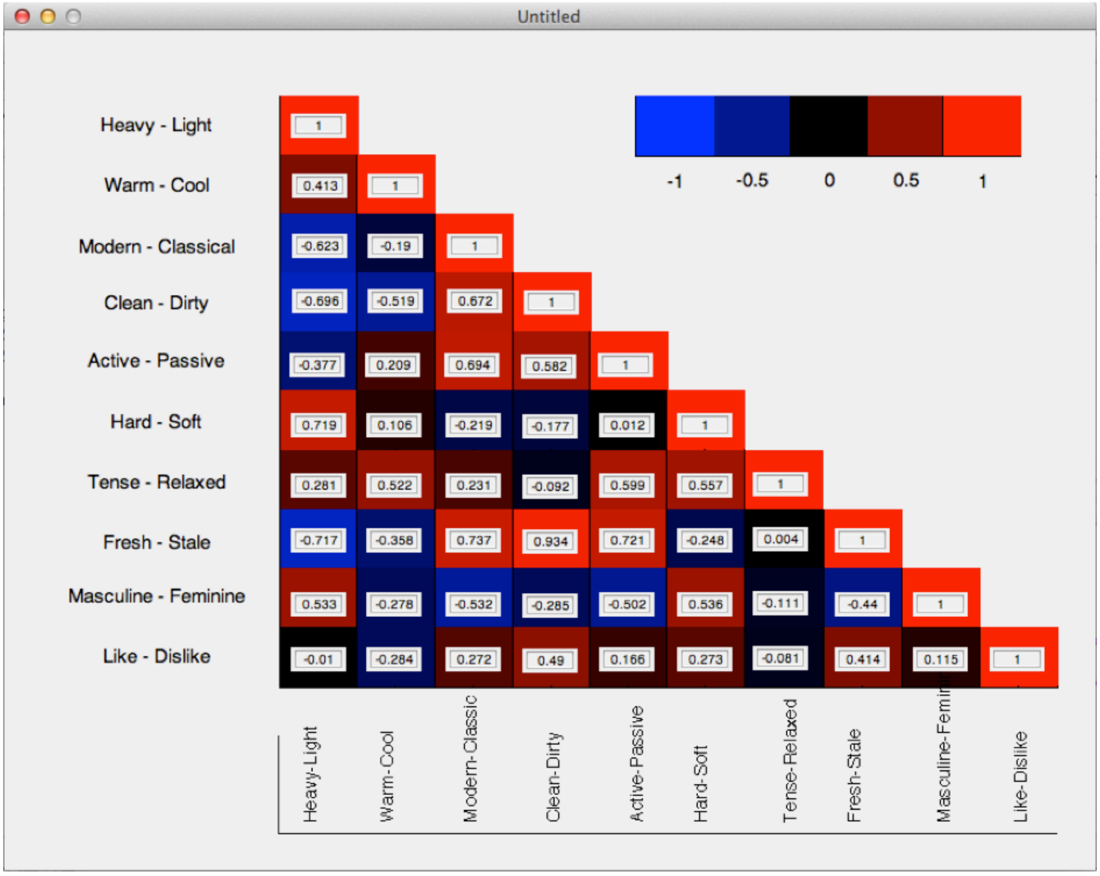


Figure 46: The correlation matrix derived from the Global Online Survey data.

7.4.3 Factor Analysis

Kaiser-Meyer-Olkin Measure of Sampling Adequacy [180, 181] and Bartlett's Test of Sphericity are used to test whether the use of principal component analysis is appropriate on the collected data or not [182]. Table 41 illustrates the KMO value of 0.66 and a Bartlett's significance level less than 0.05 indicating that the usage of principal component analysis for this data is appropriate [178].

Table 41: KMO and Bartlett's Appropriateness test for Principal component analysis.

Test	Value
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	269.721
Bartlett's Test of Sphericity Sig.	0.00

The number of principle components is identified using the Scree Plot displayed in Figure 47. Components 1, 2 and 3 seem trivial, suggesting that initially, three components may best describe the data.

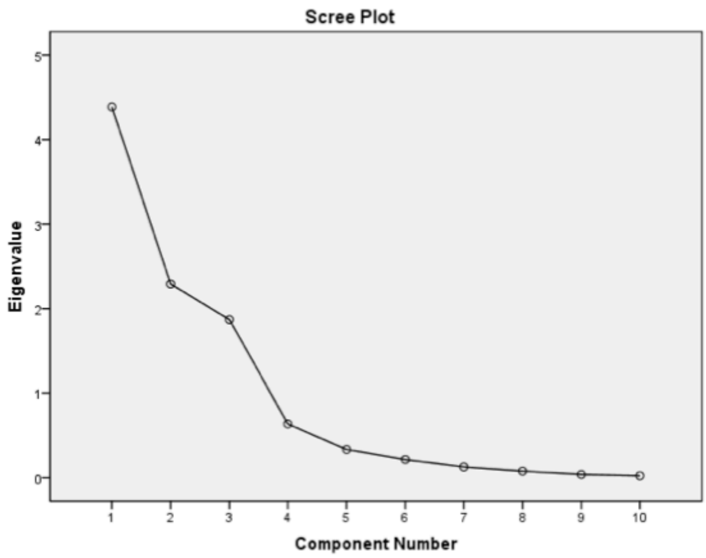


Figure 47: Scree plot for the component analysis of the Global Online Survey data.

Table 42 illustrates the rotated component matrix[85] [178]. Take the first row for example; the heavy-light loading on component 1 is -0.657, component 2 is 0.303 and component 3 is 0.620. Thus, component 3 with the highest loading of heavy-light will indicate the principal component containing other bi-polar characteristics which are alike.

---

[85] The Varimax rotation method is used to interpret the three components assuming that the factors are independent and decreases the number of variables that have high loadings on each factor.

Table 43 contains the bi-polar characteristics according to their loadings on each component. The bi-polar characteristics loaded in component 2 and 3 have positive correlation with each other. However in component 1, if a colour is rated as modern, it will also be light, clean, active and fresh

Table 42: Rotated component matrix for the three identified principal components.

	Component		
	1	2	3
Heavy-Light	-.657	.303	.620
Warm-Cool	-.364	.806	-.167
Modern-Classical	.839	.170	-.214
Clean-Dirty	.934	-.242	-.010
Active-Passive	.757	.585	-.091
Hard-Soft	-.111	.344	.893
Tense-Relaxed	.140	.896	.276
Fresh-Stale	.952	-.074	-.137
Masculine-Feminine	-.403	-.345	.716
Like-Dislike	.531	-.229	.559

Table 43: Three primary factors identified in this research (note heavy-light has a negative correlation).

Factors	Bi-polar characteristics
Component 1	Light-heavy, Modern-classic, Clean-dirty, Active-passive, Fresh-stale
Component 2	Warm-cool, Tense-relaxed
Component 3	Masculine-feminine, Like-dislike, Hard-soft

Having the correlations and also the three principal components identified, it could be concluded that the bi-polar characteristics loaded in each principal component are somehow related and react in a similar way towards colours. The next step would be to identify these relationships between colours and the bi-polar characteristics through regression models.

7.5 Fitting models to the data collected from the Global Online Survey

7.5.1 Linear models

Linear models may be less descriptive of the relationship between colour and semiotics. On the other hand, because they describe simple relationships and are easier to invert, they may be the best starting point for understanding the relationships of the data derived from the Global Online Survey[86]. First, the normality of the Z scores has been tested using the Shapiro Wilk test and have been approved. This approves the use of ANOVA in order to derive regression models. Three different linear models are derived in this section based on independent variables of L*a*b*, LCH and $LC\cos(H-50)$ which are displayed in Table 44. By a brief comparison of the models based on the L*a*b*, LCH and $LC\,cos(H°-50)$ for each of the bi-polar characteristics in Table 44, the ones with highest adjusted R^2 are chosen to be the best form of models which represent higher variance across the variables. This selection is illustrated in Table 45 and its advantage is that it offers the best form of linear models derived from the Global Online Survey data. However, one downside to this approach is that the models will not be all based on the same independent variables. On the other hand, there isn't any certain demand on having all the ten models based on the same independent variables. Therefore, this fact can be set aside for the purpose of this study. The best proposed model is for heavy-light with the adjusted R^2 of 0.86 and the weakest is proposed for like-dislike with adjusted R^2 of 0.243. It should be noted that only the proposed model for clean-dirty and fresh-stale show relevance to all three attributes of colour, whilst the rest of the eight models are based on two or one of the attributes of colour. Six of the models show significant relevance of the bi-polar characteristics with lightness (heavy-light, modern-classic, clean-dirty, hard-soft, fresh-stale, masculine-

[86] Although higher degree models such as quadratic and cubic may have more precision in describing the relationships between the dependent and independent variables they are not of interest in this thesis since there inversion methodology would not be similar to the linear models. On the other hand, although the model's performance can be estimated via the adjusted R^2 values it may not be very clear which of the models are best descriptive of the data whether quadratic, cubic or higher degree. Going back to the argument that more noise is fitted with higher degree models; it will be unclear where to stop the progression towards higher degrees. Besides, inverting the relationships is the ultimate aim which requires simplicity of the models thus, linear models would be adequate for this subject.

feminine). This shows the importance of this attribute of colour, concluding that lightness levels can be crucial in evaluating colour semiotics. A significant example is the hard-soft model, which is only based on lightness therefore colours appear softer by the increase of lightness with a coefficient of 0.015. Chroma and a* appear in 7 models with various coefficients whilst hue and b* appear in 6 models, suggesting that hue may be the less effective attribute of colour in semiotics. Looking in to this matter deeper, the coefficients of hue or b* are mostly less than 0.01. Therefore, despite its significance, it is not very effective. However, the models derived for fresh-stale and like-dislike that are based on $LCcos(H° - 50)$, show higher impact of this variable by coefficients of 0.264 and 0.206.

*Table 44: Table of linear models derived from the Global Online Survey data based on the independent variables of L*a*b*, LCH and $LC\ cos(H°)$.*

Independent variables		**Models**	**Adjusted R^2**
	1	$HL = 3.914 - 0.035L + 0.005b$	0.863
	2	$WC = 1.31 + 0.014a + 0.013b$	0.739
	3	$MC = 1.073 + 0.007L$	0.268
	4	$CD = 1.121 + 0.014L - 0.005b$	0.302
L*a*b*	5	$AP = 1.003 + 0.007a$	0.140
	6	$HS = 2.603 - 0.017L + 0.002b$	0.589
	7	$TR = 1.352 + 0.008a + 0.006b$	0.468
	8	FS =0.799+0.013L	0.223
	9	$MF = 2.858 - 0.021L - 0.09a$	0.495
	10	$LD = 1.391 - 0.11b$	0.09
	1	$HL = 4.045 - 0.033L - 0.002H$	0.822
	2	$WC = 2.672 - 0.01L - 0.004H$	0.463
	3	$MC = 0.860 + 0.007L + 0.004C$	0.301
	4	$CD = 0.591 + 0.012L + 0.006 + 0.002H$	0.441
	5	$\mathrm{AP} = 1.047 + 0.01C$	0.448
LCH	6	$HS = 2.497 - 0.015L$	0.556
	7	$TR = 1.783 - 0.007L + 0.006C - 0.002H$	0.388
	8	$FS = 0.379 + 0.013L + 0.008C$	0.438
	9	$MF = 2.285 - 0.012L$	0.185
	10	$LD = 1.283 + 0.001H$	0.111

	1	$HL = 3.903 - 0.032L - 0.004C + 0.22\cos(H - 50°)$	0.829
	2	$WC = 1.265 + 0.324\cos(H - 50°)$	0.091
	3	$MC = 0.86 + 0.007L + 0.004C$	0.301
	4	$CD = 0.924 + 0.012L + 0.008C - 0.283\cos(H - 50°)$	0.435
$\boldsymbol{LC\cos(H° - 50)}$	5	$AP = 0.680 + 0.006L + 0.01C$	0.440
	6	$HS = 2.497 - 0.015L$	0.556
	7	$TR = 1.437 - 0.005L + 0.005C$	0.207
	8	$FS = 0.319 + 0.014L + 0.009C - 0.278\cos(H - 50°)$	0.538
	9	$MF = 2.285 - 0.012L$	0.185
	10	$LD = 1.5 - 0.022\cos(H - 50°)$	0.263

Table 45: Best linear models with highest adjusted R^2 values from the Global Online survey data.

Independent Variable	**Models**	**Adjusted $\mathbf{R^2}$**
L*a*b*	$HL = 3.914 - 0.035L + 0.005b$	0.863
L*a*b*	$WC = 1.31 + 0.014a + 0.013b$	0.739
LCH	$MC = 0.860 + 0.007L + 0.004C$	0.301
LCH	$CD = 0.591 + 0.012L + 0.006 + 0.002H$	0.441
LCH	$AP = 1.047 + 0.01C$	0.448
L*a*b*	$HS = 2.603 - 0.017L + 0.002b$	0.589
L*a*b*	$TR = 1.352 + 0.008a + 0.006b$	0.468
$\boldsymbol{LC\cos(H° - 50)}$	$FS = 0.319 + 0.014L + 0.009C - 0.278\cos(H - 50°)$	0.538
L*a*b*	$MF = 2.858 - 0.021L - 0.009a$	0.495
$\boldsymbol{LC\cos(H° - 50)}$	$LD = 1.5 + 0.22\cos(H - 50°)$	0.263

7.5.2 Comparison between online survey and the lab-based experiment

Firstly it should be noted that having data collected from over 2000 participants has been helpful in terms of generating significant models for all the ten bi-polar characteristics. In comparison, Ou's research generated only 4 to 5 significant models, which were based on data collected from only 30 people. This suggests that in order to find a significant trend or relationship between colour semiotics and colour attributes, mass data collection is useful, since it allows collations of responses from diverse participants.

In order to compare the models derived from the lab-based experiment and the Global Online Survey it is useful to generate a similar table to Table 45 for the models derived in the previous chapter, which contains the best descriptive models.

Table 46: The selected linear models with highest adjusted R^2 values derived from the lab-based experiment.

Independent Variable	Models	Adjusted $\mathbf{R^2}$
L*a*b*	$HL = 3.631 - 0.04L + 0.001a + 0.005b$	0.887
$\boldsymbol{LC\cos(H^\circ - 50)}$	$WC = 2.126 - 0.009L - 0.005C + 0.91\cos(H - 50^\circ)$	0.780
$\boldsymbol{LC\cos(H^\circ - 50)}$	$MC = 0.008L + 0.007C$	0.380
$\boldsymbol{LC\cos(H^\circ - 50)}$	$CD = 0.014L + 0.013C$	0.550
LCH	$AP = 0.009L + 0.016C$	0.570
LCH	$HS = 2.965 - 0.024L + 0.006C - 0.001H$	0.710
LCH	$TR = 2.431 - 0.015L + 0.01C - 0.002H$	0.590
$\boldsymbol{LC\cos(H^\circ - 50)}$	$FS = 0.014L + 0.015C$	0.690
L*a*b*	$MF = 2.995 - 0.24L - 0.016a - 0.006b$	0.538
L*a*b*	$LD = 1.631 + 0.002b$	0.2

By calculating the difference between the two adjusted R^2 values reported in

Table 45 and Table 46, and forming a one-sample t-test, it is concluded that the mean difference is not zero. This means that there is a significant difference between the adjusted R^2 of the models derived from the lab-based experiment and the Global Online Colour Survey. The confidence interval of the mean difference is [0.031, 0.114] suggesting that 95 percent of the times the adjusted R^2 of the models derived from the lab-based data will be at least 0.03 and approximately 0.11 more than the adjusted R^2 of the Global Online Survey. Despite the significance of this difference, it is still not too high. Also, considering the fact that the Global Online Survey is based upon a considerably larger data set, generated models would be based upon a diverse data set. Thus, they would have a more generalised predictive ability rather than the models derived from the lab-based experiment. This only effects in approximately 0.11 reduction of model's descriptive ability, which can be forgivable at this point.

However, comparison of the adjusted R^2 values allows an analysis of whether the models perform equally well (in terms of the proportion of data variance) but does not test whether the models are equivalent. Thus, by taking 1000 random L*a*b* values and computing the bi-polar characteristics for both sets of models, the correlations between two separate models can be identified (see Figure 48). The best linear models produced with highest adjusted R^2 values are chosen to be compared with each other. Interestingly, the two models are highly positively correlated with each other and can be descriptive of the relationship between the colour space and semiotics in the same way. The like-dislike model stands out from the rest by having the least correlation coefficient, which could be a general result of the model's weakness in describing the relationship. Over all it can be concluded that the models derived from both web-based and lab-based experiment are well correlated and have the ability to describe and predict in almost the same manner.

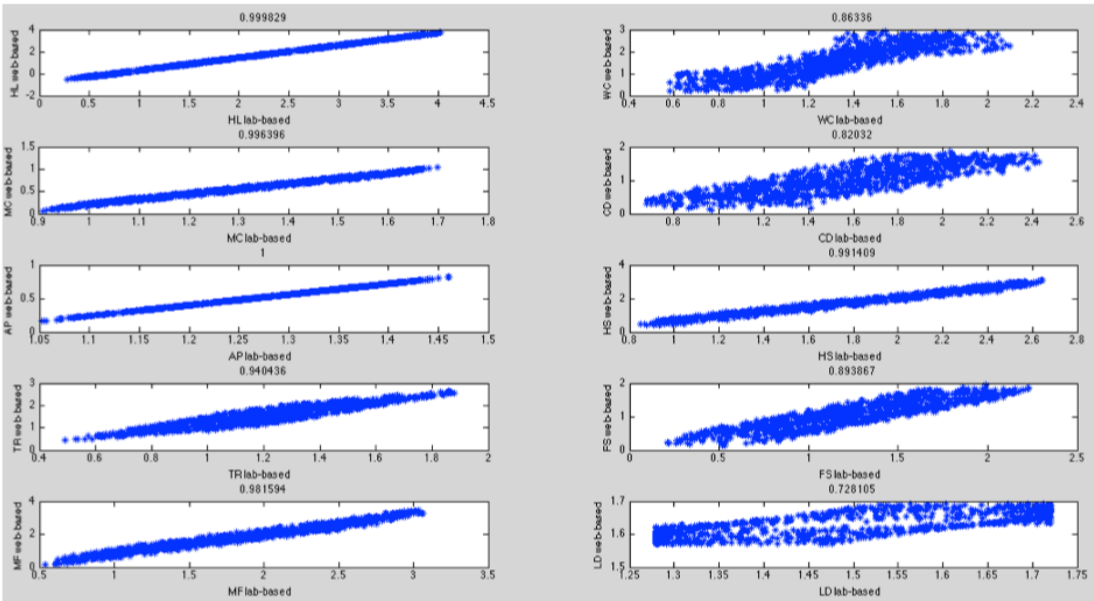


Figure 48: Scatter diagrams of the 1000 simulated colours and there bi-polar characteristics, derived from lab-based and web-based experiments.

7.5.2.1 Inversion of the linear models

The ultimate application of the colour semiotic models is based upon the inverse relationship of colour and semiotics. Regardless of the specific independent variables used in the chosen models (whether it is L*a*b*, LCH or $LC\cos(H-50)$), the gradual approximation technique used in the previous chapter is applicable for the linear models derived from the Global Online Survey. These models are successfully applied to tool such as illustrated in .

7.6 Effective personal variables

In search of an effective colour selection tool for design process, it is important to consider most of the effective variables that designers consider whilst focusing on the target market such as age, gender and culture of the market. For example, a designer would like to choose a colour for a hand bag which is designed for European females in their 20s; it may be different from a colour chosen for Asian males in their 50's. Consideration of these variables in the tool requires identification of significant models which contain these variables. One of the main things which makes this research different from the previous studies is the unique nature of the gathered data giving, the opportunity to investigate each of the effective variables such as gender, age, culture and etc. In this section, the significance of these variables is evaluated based on the gathered data using ANOVA and t-test. Bearing in mind that the data should be normally distributed with similar sample size and independence amongst the groups in order to fit the assumptions of these tests. In other words, in order to enter the gender parameter into the models which describe the relationship between colour and semiotics, it would be necessary to split the data set into categories that define these variables such as male and female. Therefore a new set of Z scores is calculated for each category followed by relevant investigation of ANOVA and correlation. Initially, separate models are derived for each of the gender groups. Note that the modelling procedure would be exactly the same as applied for derived data from the lab and also the Global Online Survey, along with dummy variables would be added which identify the effective personal variables.

7.6.1 Gender

Descriptive reports of the Global Online Survey show that females have taken part almost 19% more than men with a participation of 1356. We assume that this un-equality will not disturb the analysis of Z scores. However it is interesting to conclude that overall, females show more interest in taking part in colour related surveys. In order to find out whether gender can be considered as a significant factor, the Z scores derived from females and males are compared.

7.6.1.1 Results

The correlation between the Z scores derived from female and male data is plotted for the ten bi-polar characteristics in Figure 49. The reported correlation coefficients are all positive and higher than 0.65 suggesting that the male and female participants have been responding in a very similar way and there may be little difference between them. In order to find out whether this difference is significant a one-way ANOVA with p-value<0.05 is applied which results that the bi-polar characteristics are the same across the gender categories.

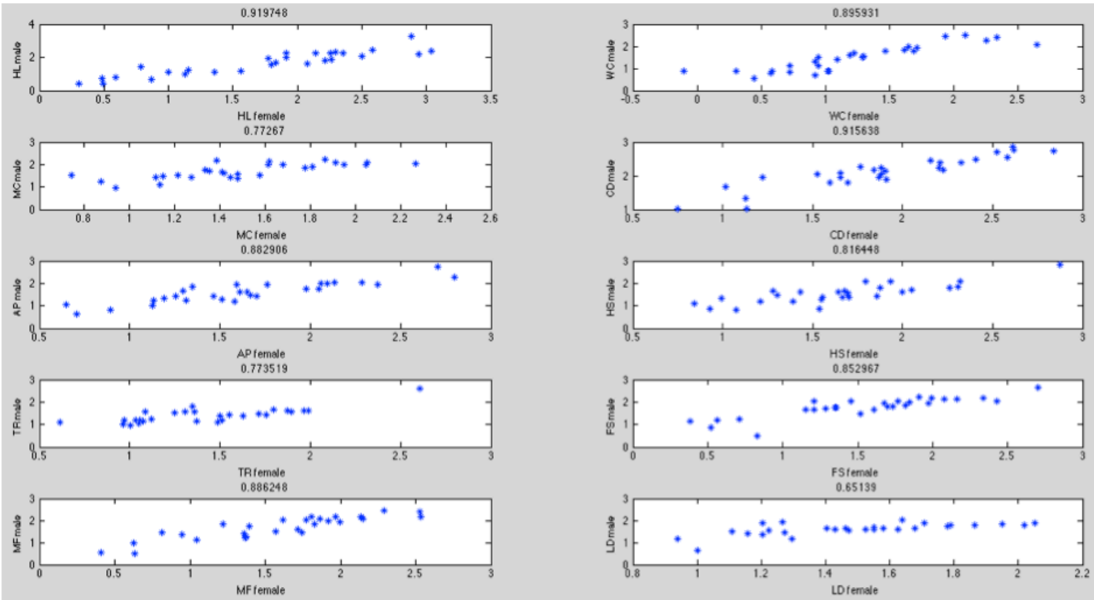


Figure 49: Correlation plot of the ten bi-polar characteristics, for Z scores derived from male (y axis) and female (x axis) data (correlation coefficients are displayed above each figure).

High positive correlation between the two gender groups and also the results of the Kruksal Wallis test strongly suggest that no significant difference can be remarked between males and females medians. However, further investigation is made for modelling the relationship between colours and semiotics by adding the gender parameter as a dummy variable. Table 47 illustrates models that show significant coefficients for gender parameter with highest adjusted R^2. Note that these models are chosen from the significant models based on the independent variables of L*a*b*, LCH and LCcos$(H-50)$. The variable *G* indicates gender, which adopts the value 1 for female and 0 for male participants. Warm-cool, Modern-classic, clean-dirty and fresh-stale show significance towards gender with a coefficient of almost -0.21. This means that generally females tend to perceive colours 0.21 cooler, classical, dirtier and stale compared to males.

Table 47: Linear models with significant gender effect. Note that G=1 for female and G=0 for male.

Model	Adjusted R^2
$\boldsymbol{WC = 1.466 - 0.231G + 0.014a + 0.013b}$	0.726
$\boldsymbol{MC = 1.063 - 0.209G + 0.007L + 0.004C}$	0.318
$\boldsymbol{CD = 0.756 - 0.213G + 0.012L + 0.005C + 0.002H}$	0.455
$\boldsymbol{FS = 0.605 - 0.227G + 0.013L + 0.008C - 0.293\cos(H - 50)}$	0.522

7.6.1.2 Conclusion

According to the correlations, Kruksal Wallis test and the un-successful attempt to modelling, dark brown seems to be heavy, no matter if it is asked from a woman or a man and also the same for active-passive, hard-soft, tense-relaxed, masculine-feminine and like-dislike. However, warm-cool, modern-classic, clean-dirty and fresh-stale show significance to the dummy variable of gender, indicating that there is a difference between male and female response for four of the bi-polar characteristics. Previous research such as Ou's has failed to show this significance, which can be a result of small sample size. Ou used 30 participants in his experiments[87] whilst in the Global Online Survey 1356 females and 916 males have taken part, which helps to identify its effectiveness.

7.6.2 Age

Descriptive reports show that the oldest person taking part in the survey is 78 while youngest is 18 years old stating the average age of participants as 48. In order to identify the effectiveness of age, it is preferred to categorise participants into certain age groups so that more data would be available for analysis. The boundaries of the age groups are completely empirical, since it would vary in accordance to the application of models to a specific design. For example, a designer may wish to use the colour semiotic models for the

[87] Number of males and female are not indicated.

age range of 25-30 and another is interested in only 35 years olds. In the future, it may be possible to generate models for each of the required age groups upon request. For the purpose of this research age groups are categorised in to 4 categories of [18,32], [33,47], [48,62] and [63,78].

7.6.2.1 Results

The frequencies of the four age categories are displayed in Table 48. Individual Z scores are calculated for each of the four groups and are tested for normality using the Shapiro Wilk and Bartlett's test for equality of variance in order for ANOVA to be applied and linearly modelled. The results of the ANOVA test show that the active-passive, tense-relaxed and fresh-stale are different across the four age groups; Tukey's test is used to identify which of the age groups are different from each other in these characteristics. Age groups 1 and 3 respond differently to active-passive, hard-soft, tense-relaxed and fresh-stale. Also age groups 1 and 2 along with 1 and 4 respond differently to tense-relaxed, so do age groups 3 and 4 to active-passive. This means that in modelling the relationship between colour and semiotics for these bi-polar characteristics, certain age group variables should be considered. The Pearson's correlation coefficient is calculated for each of the bi-polar characteristics and the four age groups (see Table 49). Over all age groups 2 and 4 and also 3 and 4 are not well correlated across the ten bi-polar characteristics showing the difference response between younger and older people. Table 50 illustrates the models found to be significant which contain age group dummy variables. In these models, the value of the variable Age1 is equal to 1 for respondents of 18 to 33 years old and 0 for the rest. Note that these models have been chosen from the best models developed from independent colour variable of L*a*b*, LCH and $LCcos(H-50)$.

Surprisingly, active-passive and masculine-feminine contain two age group dummy variables 2 and 3. This is subject to the high Pearson' s correlation value in Table 49.

Table 48: Table of age groups and frequencies

	Age range	Frequency	Relative Frequency
1	[18,32]	1252	0.55
2	[33,47]	606	0.27
3	[48,62]	322	0.14
4	[63,78]	93	0.04

Table 49: Table of Pearson's correlations between age groups for the ten bi-polar characteristics.

	Age group correlations					
	1,2	1,3	1,4	2,3	2,4	4,3
HL	0.97	0.90	0.75	0.86	0.74	0.71
WC	0.88	0.83	0.50	0.86	0.51	0.41
MC	0.68	0.60	0.12	0.59	-0.07	-0.03
CD	0.86	0.79	0.45	0.82	0.69	0.64
AP	0.89	0.74	0.57	0.75	0.59	0.44
HS	0.80	0.54	0.59	0.57	0.61	0.30
TR	0.84	0.64	0.71	0.54	0.59	0.49
FS	0.87	0.85	0.27	0.85	0.43	0.32
MF	0.89	0.85	0.14	0.78	0.06	0.11
LD	0.61	0.51	0.41	0.36	0.25	0.37

Additionally it could be added that masculine-feminine is the only model, which shows significant relationship to all age groups 1, 2, 3 and 4 (age group 4 is referred when all the three dummy variables are equal to zero).

Table 50: Linear models with significant age effect. Note that Age1=1 for people age group 1 and 0 for the rest.

Models	Adjusted R^2
$HL = 3.699 - 0.035L - 0.003a + 0.003b + 0.223Age2$	0.723
$AP = 0.795 + 0.011L + 0.007a + 0.004b + 0.253Age2 + 0.553Age3$	0.326
$HS = 2.691 - 0.016L + 0.002b - 0.283Age1$	0.441
$TR = 1.873 - 0.004L + 0.006a + 0.006b - 0.352Age1$	0.428
$FS = 1.002 + 0.012L - 0.002B - 0.285Age1$	0.252
$MF = 2.489 - 0.018L - 0.008a + 0.264Age1 + 0.269Age2 + 0.33Age3$	0.341
$LD = 1.639 - 0.166Age1$	0.030

However, at this point a very important issue should be considered which is the un-equality of the frequency values of the 4 groups. The calculation of Z scores is influenced by the number of participants as the following holds; $z - 1.96\frac{\sigma}{\sqrt{n}} \leq E[z] \leq z + 1.96\frac{\sigma}{\sqrt{n}}$. Therefore; it is necessary to weight the Z score by the following equation;

Equation 72

$$w_i = \frac{f_i}{2273}$$

where f_i indicates the number of respondents in a certain group and w_i is the weight which the Z scores would be multiplied by.

By this, the high values of the Z scores, which belong to the groups with smaller frequencies, would become less effective compared to the Z scores derived from groups with higher frequencies.

Table 51: Table of weighted linear regression coefficients of dummy variables for age categories (multiply the values by 10^{-3}).

	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
Constant	591	55	-109	-157	-88	327	67	-158	380	126
L*	-9	0	3	4	2	-4	0	4	-5	-1
a*	0	4	1	0	2	0	2	0	-2	0
b*	1	3	0	-1	1	0	1	-1	0	0
C	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
$\boldsymbol{Cos(H-50°)}$	0	0	0	0	0	0	0	0	0	0
Age 1	832	668	783	1027	734	724	631	729	844	742
Age 2	416	299	418	438	409	409	362	394	388	369
Age 3	162	137	175	211	225	195	156	201	183	156
Adjusted $\boldsymbol{R^2}$	829	784	865	876	745	850	830	754	817	899

One of the advantages of using the weighted Z scores for the linear regression is that more significant models have been developed for the ten bi-polar characteristics with considerably high adjusted R^2 values. Overall age group 1 has higher coefficients compared to age groups 2 and 3. Also, seems that considering age groups as dummy variables remarks L*a*b* as significant attributes of colour rather than LCH or $LCcos(H-50)$.

7.6.2.2 Conclusion

According to the defined categories, there are significant differences between how age groups respond to colour semiotics. Although the respondents who belong to the younger age groups (groups 1 and 2) are more likely to respond in the same way (see columns 1 to 2 of Table 49), correlations become less with older age groups. However, correlation does not report the effectiveness of the individual age groups and therefore seven significant models have been identified containing the age group dummy variables. Leaving warm-cool, modern-classic and clean-dirty as the only ones that do not show significant relationship with age groups. It should be reminded that the identified relationships may vary by a different categorisation of the age group. Also, since the number of participants has not been very high in age groups 3 & 4, and also colours have been presented randomly, it has not been possible to collect responses on all the 28 colours. Overall it can be concluded that age is an important factor and should be considered when it comes to defining the relationships between colour and semiotics. This means that a designer is required to consider the target age category prior to colour selection. The application of these models in the colour selector tool is in a way that the user/designer selects a specific age category and the relative models would be applied upon request. For age categories that show no significance; the normal models would be applied.

7.6.3 Culture and nationalities

Regarding to the question in the survey about participant's chosen language, native language and nationality; there are several different way to identify the effect of culture. Figure 50 displays the percentage of the chosen languages indicating that English has been chosen more than 60 percent of the time followed by Persian language for 13.42 percent of the time. Italian and French have been chosen 6 and 5 percent of the times while Malay and Urdu language have the least percentage (less than 1.32 percent of the times). Having English chosen so many times may not necessarily mean that all were native English speakers, since participants with native languages other than the specified ten languages were forced to choose this option[88]. This is why, all participants were asked to manually enter their native language and nationality on the second page so that later on it would be traceable (see Figure 53). Overall, 67 different nationalities have taken part in the survey. Having detailed information about participant's native language and nationality helps tracing the cultural effects in colour semiotics.

[88] This could be because they may not have their native language installed on their Internet browser forcing them to respond in English or that they feel more comfortable in responding in English.

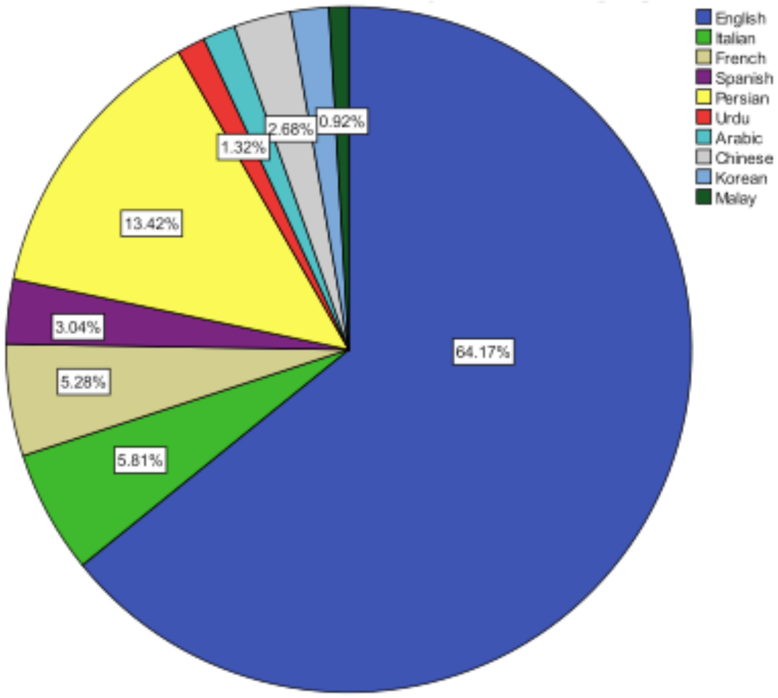


Figure 50: Pie chart of participant's chosen language percentage.

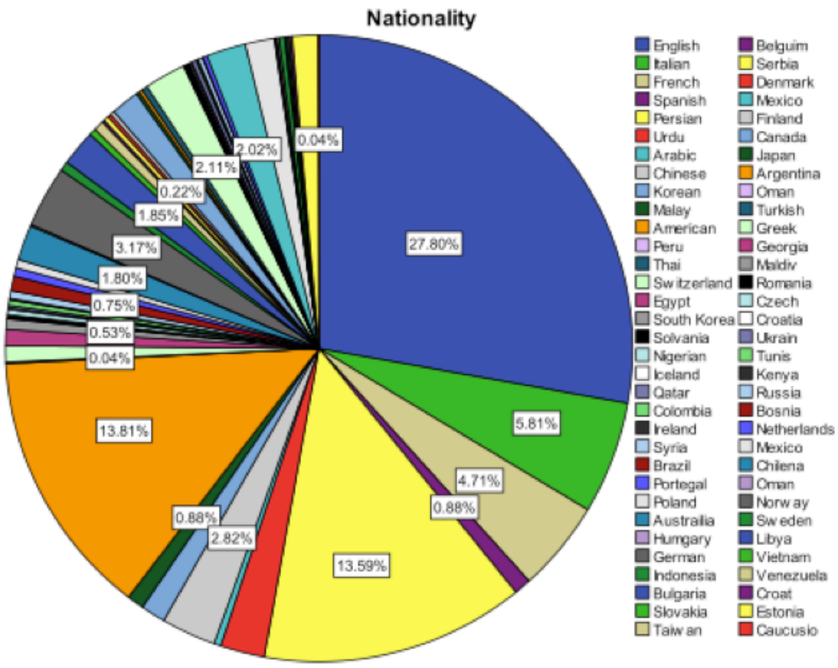


Figure 51: Pie chart of the participant's nationality and the percentage.

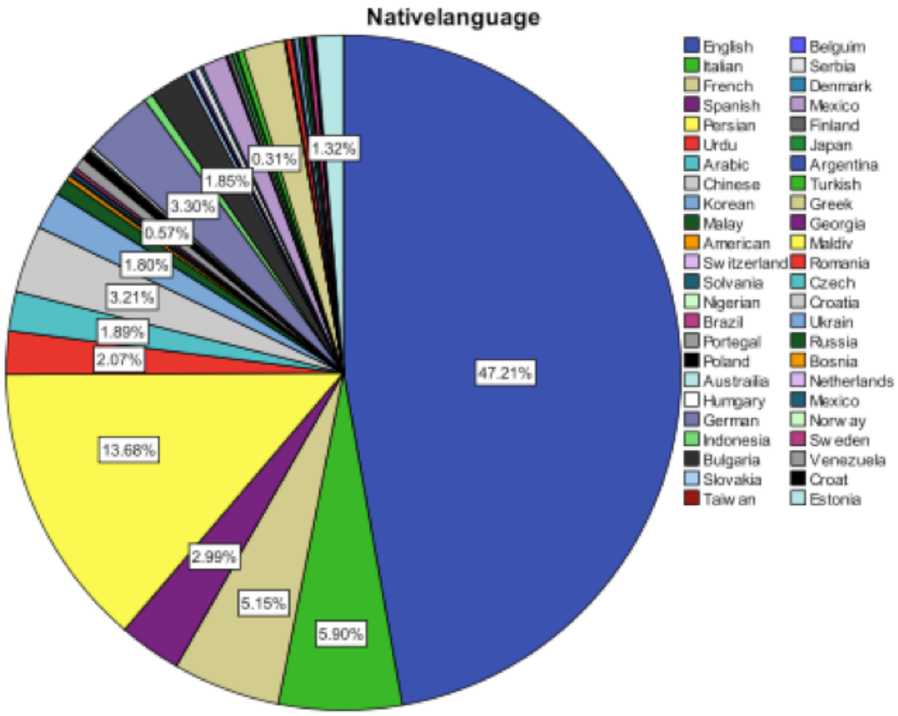


Figure 52: the participant's native language percentages

7.6.3.1 Results

The relationship between colours and bi-polar characteristics has been modelled by considering the chosen language as dummy variables[89]. For example; if Spanish is chosen, then the relevant dummy variable to Spanish would be equal to 1 and all the other dummy variables will be zero. A primary analysis of the effect of language shows relatively high impact. Arabic language has shown the most significance for 8 of the bi-polar characteristics followed by Persian, Chinese and Urdu with significance for 6. On the other hand, Italian and Spanish show significance for only two and English and Korean for three of the bi-polar characteristics leaving French significant for four of the bi-polar characteristics. Overall, it can be concluded that eastern languages show more significance compared to the western (with the exception of Korean). However, it is very important to note that many of the participants whom chose English were not actually British but American or

[89] Considering the participation of high number of cultural groups, it would be very hard and time consuming to find all the two by two correlations between the groups, or to even find the differences between them. Considering them as dummy variables and testing the significance of the coefficients can be considered sufficient for the purpose of this section.

Australian and even not western (such as Indian, Japanese and etc.). This huge variety of response amongst the data collected from English language makes it harder to find a significant relationship. This is why it would be useful to use the collected data from the question regarding to the native language, in order to split the data for English language. This would be helpful in the sense that more information will be used. On the other hand, if all the data collected from the nationality question would be used then there would be too many categories with not much representative data in each. Thus, it will only be used to distinguish the categories in data gathered from English language. Figure 53 (left) displays the variety of different participant's nationalities that chose to take part in the survey with English language.

Table 52: Table of linear regression coefficients, with language dummy variables (multiply the values by 10^{-3}).

	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
Constant	3274	1695	948	1060	652	246	1893	400	2274	1442
L*	-24	-4	5	8	5	-13	-6	110	-13	0
a*	0	9	0	0	0	0	0	0	-8	0
b*	4	10	0	0	0	2	0	0	-2	0
C	0	0	4	5	10	0	4	9	0	0
H	0	0	0	1	-1	0	-1	0	0	0
$\boldsymbol{Cos(H-50°)}$	0	0	-163	0	0	0	0	-282	0	-197
English	0	0	0	0	409	0	0	257	223	0
Italian	-314	0	0	-327	0	0	0	0	0	0
French	-343	-250	0	-459	267	0	0	0	0	0
Spanish	0	-330	0	-408	0	0	0	0	0	0
Persian	-339	-278	0	-334	0	-496	-748	0	0	404
Urdu	-497	0	0	-628	-398	-494	-444	0	-554	0
Arabic	-399	0	-547	-755	-451	-663	-696	-739	-582	0
Chinese	-387	-262	0	-633	0	0	0	-270	-283	-314
Korean	0	0	0	-619	357	0	0	-472	0	0
Malay	-1083	-787	-605	-1350	-340	-1032	-696	-462	-782	-463
Adjusted R^2	435	344	151	327	310	332	295	331	284	152

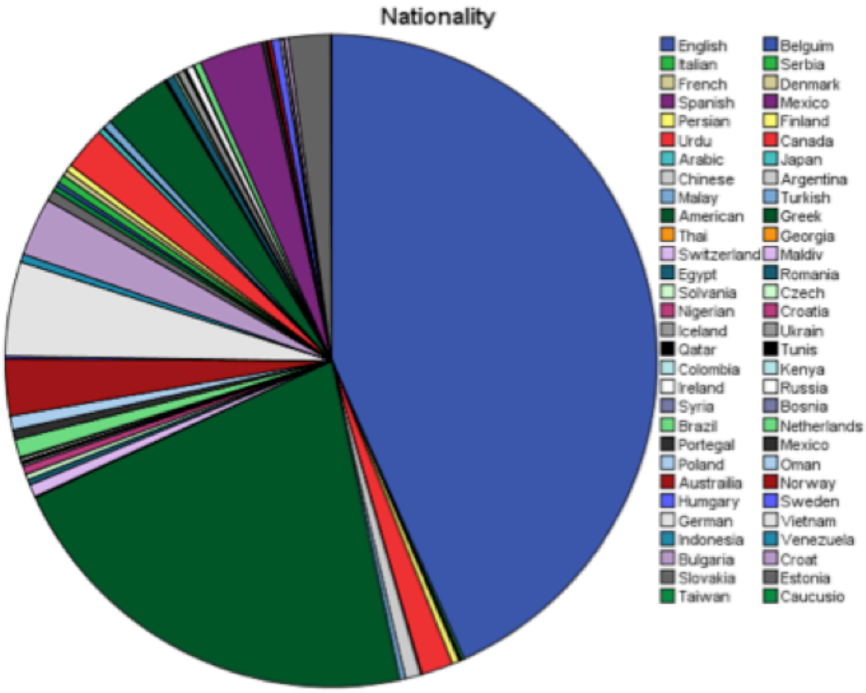


Figure 53: Pie chart of the different nationalities whom chose to take part in English language.

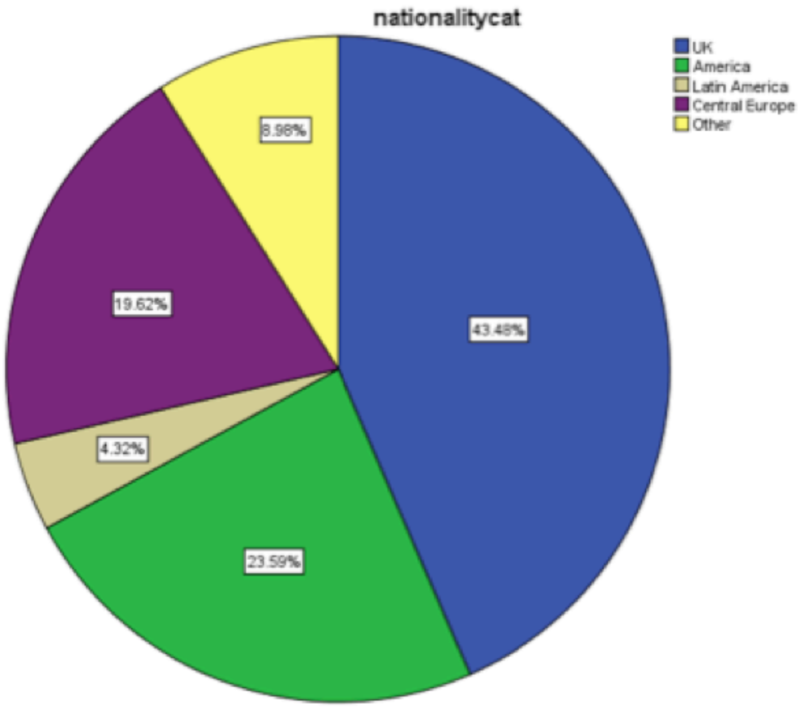


Figure 54: Pie chart of the nationalities categories.

This suggests that this data can be categorised as illustrated in Figure 53 (right) where America indicates North American and Canadian nationalities. Latin America indicates Mexican, Brazilian, Colombia and Peru. Central European indicates Germany Greece, Serbia, Hungary, Georgia, Bosnia, Switzerland and etc. and other indicates Russia, Taiwan, Indonesia, Japan, Egypt, Qatar, Syria, Turkey, Australia, Tunis etc. Resulting in 15 groups of Italian, French, Spanish, Persian, Urdu, Arabic, Chinese, Korean, Malay, UK, America, Latin America, Central Europe and other. By this categorisation, the ten language categories are extended to 14. Additional Z scores are calculated for the 5 new categories and combined with the 9 previous Z scores. Note that the effect of the 14 th category (named as other) is added to the constant.

However, the un-equality of the frequency values of the 14 groups requires the weightings of the Z scores. By this, the high values of the Z scores, which belong to the groups with smaller frequencies, would become less effective compared to the Z scores derived from groups with higher frequencies. However, it is thought that this will change the structure of the linear relationship which possibly could be the reason behind the rise of the adjusted R^2 value reported in Table 54**.**

Table 53: Table of linear regression coefficients, with the 13 dummy variables (multiply the values by 10^{-3}).

	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
Constant	3097	1465	892	683	720	2444	1565	540	2413	1542
L*	-26	-4	6	9	4	-14	-4	11	-15	0
a*	0	10	0	0	0	0	5	0	-8	0
b*	4	11	0	0	0	2	4	0	-2	0
C	0	0	5	5	10	0	0	8	0	0
H	0	0	0	2	-1	0	0	0	0	0
$\boldsymbol{Cos}(\boldsymbol{H}-\mathbf{50}°)$	0	0	-146	0	0	0	0	-275	0	-186
Italian	0	0	0	0	0	0	0	0	0	0
French	0	0	0	0	0	0	361	0	0	-263
Spanish	348	0	0	0	0	0	259	0	0	0
Persian	0	0	0	0	0	-456	-534	0	0	302
Urdu	-189	352	0	-297	-433	-454	-230	0	-537	0
Arabic	0	0	-587	-424	-487	-622	-483	-866	-565	-211
Chinese	0	0	0	-302	0	0	0	-397	-267	-416
Korean	0	0	0	-288	322	0	341	-599	0	0
Malay	-775	-544	-645	-1020	-376	-992	-483	-589	-765	-565
Uk	0	0	0	0	274	0	399	0	354	0
America	273	0	0	255	538	304	379	263	318	0
Latin America	0	0	-248	-246	0	0	0	0	0	-257
Central Europe	0	334	0	0	315	0	0	0	0	0
Adjusted R^2	489	410	182	323	331	348	316	363	325	155

Table 54: Table of weighted linear regression coefficients of 13 dummy variables (multiply the values by 10^{-3}).

	HL	WC	MC	CD	AP	HS	TR	FS	MF	LD
Constant	416	120	33	40	61	280	178	-1	303	157
L*	-4	0	1	1	1	-2	0	2	-2	0
a*	0	2	0	0	0	0	1	0	-1	0
b*	1	2	0	0	0	0	1	0	0	0
C	0	0	0	1	1	0	0	1	0	0
H	0	0	0	0	0	0	0	0	0	0
$\boldsymbol{Cos(H-50°)}$	0	0	0	-36	0	0	0	-39	0	-28
Italian	0	0	0	0	0	0	0	0	0	0
French	0	0	0	0	0	0	0	0	0	-35
Spanish	0	-59	-80	-79	-77	-73	-63	-67	-60	-70
Persian	197	159	221	250	185	117	45	174	177	283
Urdu	-116	-83	-115	-135	-126	-131	-124	-123	-127	-119
Arabic	-109	-87	-124	-133	-124	-131	-126	-135	-124	-117
Chinese	-80	-63	-81	-100	-88	-86	-85	-98	-91	-100
Korean	-98	-75	-105	-122	-93	-104	-95	-123	-103	-106
Malay	-135	-109	-134	-158	-131	-148	-136	-139	-137	-138
Uk	650	571	709	781	668	687	719	684	759	651
America	323	223	287	374	367	357	316	345	334	284
Latin America	-71	-55	-87	-95	-78	-80	-88	-79	-81	-91
Central Europe	177	225	223	253	230	219	197	215	192	183
Adjusted R^2	749	772	882	874	853	870	885	843	852	915

In order to compare the magnitude of the effectiveness of each of the cultural groups, coefficients of the derived models are plotted for each of the ten bi-piolar characteristics in Figure 55. Overall, it can be concluded that the effect of the main 13 cultural groups are the same across the ten bi-polar characteristics, apart from French and Italian (refers to small number of participants in this group). This means that all the ten bi-polar characteristics are affected approximately in the same way when presented to a certain cultural group. More importantly, it is concluded that the cultural effects are significantly different from one another for each of the ten bi-polar characteristics. This is evident by the sudden rise and fall of the lines for each of the cultural groups. Specifically, the peaks of this fluctuation happen on Persian and UK where a lot of response has been gathered. On one hand; this fact may point to the effect of different frequencies, suggesting that even weighting the Z scores is un-sufficient.

But in general, no matter how the weighting is applied, the Zscore signs will remain the same. By taking a look at the graph it is evident that coefficients are always negative with Urdu, Arabic, Chinese, Korean, Malay and Latin American groups and positive for Persian, UK, America and Central Europe.

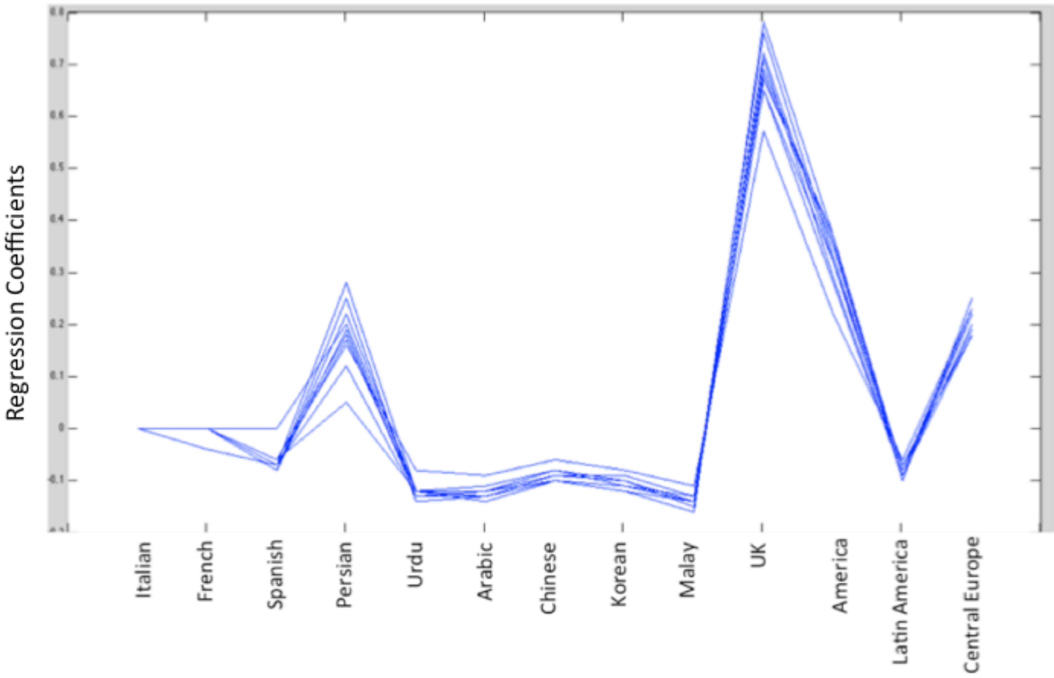


Figure 55: Plot of the linear regression coefficients of each of the cultural groups. Lines indicate the ten bi-polar characteristics.

In order to find out which of these groups act in the same way, factor analysis is carried out on the weighted Z scores. The results of the factor analysis indicate that cultural groups can be factorised into three principal components. According to the Varimax rotated component matrix illustrated in Table 55, the cultural groups can be factorised as Table 56. By this factorisation, it can be concluded that Urdu, Arabic, Chinese, Korean, Malay and Latin-America respond to colour in the same way whilst UK and America respond with the exact opposite. A similar analysis holds for Spanish versus Central Europe. Persian falls in to one component on its own and does not show any similarity to other groups. It should be noted that these analysis have been based on the coefficients of the regression models rather than the pure Z scores. Also, the Italian and French coefficients were zero in many cases and therefore were not used in the factor analysis.

Table 55: Varimax rotated component matrix.

	Component		
	1	2	3
Spanish	.211	.880	.103
Persian	-.033	-.040	-.956
Urdu	.979	.032	-.045
Arabic	.938	.228	-.054
Chinese	.840	.035	.427
Korean	.819	.106	.390
Malay	.928	.079	.147
Uk	-.851	-.186	.029
America	-.867	-.086	.178
Latin America	.817	.170	.298
Central Europe	-.053	-.909	.029

Table 56: Loadings of the cultural groups on the principal components.

	Positive	Negative
1	Urdu-Arabic-Chinese-Korean-Malay-Latin America	Uk-America
2	Spanish	Central Europe
3		Persian

7.6.3.2 Conclusion

In this section, the effect of the cultural groups has been analysed. The challenges faced at this part of the analysis rises with the in-equality of the cultural group sizes. This is overcome by the multiplication of proportional weights in the Z scores. Linear regressions are derived for each of the ten bi-polar characteristics with 13 dummy variables indicating 14 cultural groups. The coefficients show the amount of significance of each group and therefore tested by the factor analysis. The results show that the response to colour from Urdu, Arabic, Chinese, Korean, Malay and Latin America languages is the opposite of UK and Americans. The very same holds for Spanish versus Central Europe, whilst Persian responses fall in to one separate component. In comparison to age and gender, it seems that culture is more effectively in relation to colour and must be considered as a very important parameter in colour semiotics.

7.7 Discussion

In this chapter, the data collected from the Global Online Survey has been analysed and modelled. Linear models have been of special interest since they illustrate the relationship between effective variables and colour semiotics in a simplicity, such that later on it could be inverted using the approximation technique discussed in the previous chapter. In the beginning of this chapter, investigation for the best linear models which describe the relationship between colour attributes (L*a*b*, LCH or $LCcos(H-50)$) and semiotics has taken place on the whole data collected from the Global Online Survey all together. The model performance metric of adjusted R^2 has been used for the selection of the best models. These models are then compared with the models derived from the Lab-based experiment. The effectiveness of additional variables such as age, gender and culture have also been investigated thoroughly and modelled. It seems that gender differences only have considerable influence on four bi-polar characteristics whilst on the other hand, differences in age and culture are more broad. However, this depends on the number of participants in each group and also the categorisation boundaries. Therefore, it is completely relevant to the structure of the variables of interest. Since the different series of models derived in this chapter are based on different type of variables such as age, gender and culture it would be meaningless to compare the adjusted R^2 values. However it can be concluded that although by adding these variables to the models, the overall performance of the model may decrease, it is still beneficial since it assumes other crucial psychophysical variables of colour as well as the physical. Overall, it is clear that culture is a very effective parameter and shows significant variance between different groups; Gender on the other hand does not show that much difference. Also, by an overlook of all the models derived for the ten bi-polar characteristics; it is evident that like-dislike always acts differently from the other nine bi-polar characteristics. With regard to the definition of colour semiotics, it is more likely to categorise like-dislike under colour preference rather than semiotics. This enlightens the fact that colour preference is different from colour semiotics and may not be modelled in the same way.

Overall, the findings of this chapter conclude that it is possible to find significant mathematical models, which can describe meaningful trends between colour semiotics, and physical and psychophysical aspects of colour. This opens a whole new door to the understanding of colour and its relativity to different variables.

Additionally by this chapter, the proposed methodology of the online data collection is proved to be effective and that it provides similar response to the ones derived from the Lab-based experiment. This highlights that mass data collection not only would hide the effect of un-controlled conditions and huge variation of viewing conditions but provides more reliable and realistic data that is freshly collected from all around the globe. Considering categorical-judgment scales rather than the common paired comparison method, also gives a huge lift to the amount of information in hand, and, together with the novelty of online data collection, forms a unique insight in to our understanding to colour semiotics.

More and more data are to be collected from the survey each day, which keeps this experimental updated. Eventually there will be thousands of participants that would open doors for identifying new undiscovered effective factors in the future.

8 Final Discussions and Implementations

8.1 Introduction

In this chapter the entire research carried out throughout this thesis is summarised and contributions are addressed. Also, despite the effort made towards modelling the relationships, the complex mathematical appearance and structure of these models may not be ideal for certain applications. Note that the high number of models generated in this thesis for each of the parameters makes it even harder for the user to apply each model in its right place. Getting back to one of the main motivations of this research, which is to bridge the gap between design and colour science, it is important to investigate an implementation or in other words more comprehensive application from these models which offers the models in a more user-friendly environment. Therefore at the end of this chapter the implementations of the results of this research are enlightened and also possible future work that can be carried out are noted.

8.2 Summary

This thesis is concerned with the relationship between colour and certain characteristics. In this work bi-polar characteristics have been used; thus, whether, for example, a colour is warm or cool. This topic, variously referred to in the literature as colour emotion or colour semiotics or even colour language has been studied quite intensely over the last decade or so. Probably the work of Ou and colleagues [1] can be considered to be the leading research in this field. Ou *et al.* carried out psychophysical experiments and developed models (by fitting the psychophysical data with parametric equations) that predict bi-polar characteristics from colour values. In Chapters 2, 3 and 4 thorough review of the relevant literature has taken place. However, the research in this thesis starts from the question of whether experiments conducted in the laboratory (which are necessarily constrained to have low numbers of participants) can ever hope to capture the full complexity of the relationships being studied since there are likely to be strong cultural and regional differences. The key advance in this work therefore was to explore the use of a web-based experiment for collecting data on a large scale and from all over the world. The implementation of a web-based experiment in colour is subject to possible criticism. However, the colours displayed (and the viewing environment) will not be as well controlled as in a laboratory experiment and it is not easy to screen-out colour-defective observers. Nevertheless, there are examples of successful web-based colour experiments, despite the problems and the work in this thesis was in part inspired by the work by

Moroney who has been carrying out a web-based colour-naming experiment for at least 6 years [2].

In Chapter 5 of the thesis a laboratory-version of the web-based experiment was carried out. This was to enable a comparison between laboratory- and web-based paradigms and also to be able to relate the work in this thesis with the seminal work of Ou and colleagues [1]. The data from the experiment show some correlation with values derived from Ou's models. However, the correlation values for some of the characteristics such as active-passive are not very high. There were some differences between the experimental details of the work in this chapter and those carried out by Ou. For example, there were almost 30 observers in Ou's experiment and 20 in this experiment. Also, there were 20 colours in Ou's experiment and 28 in this experiment (and the colours were not the same). Most importantly, Ou used a paired-comparison method whereas in this work a categorical-judgement method was used. In general it may be concluded that despite the differences in methodology the derived models are broadly similar. In some ways this would seem to suggest that a laboratory experiment with relatively few participants is acceptable and robust; however, such an experiment is likely to be very limited in terms of the variation in culture that exists globally and the advantage of a large-scale web-based experiment is that variables such as age and cultural differences could be explored.

Some preliminary modelling of the data in Chapter 5 was carried out. However, in Chapter 6 the main aim was to investigate how to address the inverse problem for models that predict bi-polar characteristics from colour attributes. For almost all practical applications of these models a user will start with a set of desired bi-polar characteristics and will want to know which colours produced them. Firstly, linear models based on L*, a* and b* values were derived for each of the bi-polar characteristics. Although these models were less accurate than non-linear models based on LCH values, they were selected for the application since they could be inverted more easily. Note that, four of the bi-polar characteristics were modelled with relatively high R^2 values. In general it was concluded that only one or two of the models were accurate. Therefore using simple models to allow invertibility may result in too much loss of accuracy. This is a trade-off between accuracy and invertibility and unfortunately the terms of the trade-off are probably not acceptable. Secondly, a numerical method (adaptive gradient descent optimisation) was implemented to solve the inverse problem. The advantage of this approach is that it can work with linear or nonlinear models. In this chapter we developed the optimisation method using linear models. It was necessary to develop an sRGB gamut mapping algorithm to ensure that solutions were physically reasonable.

The global online colour survey was described in Chapter 7. Data collected from the Global Online Survey was analysed and modelled. Data were collected for over 2000 participants from 67 countries in one of 10 languages. In order to encourage high participation rates a novel experimental paradigm was implemented whereby each participant only considered one colour; the experiment was therefore distributed over a large number of observers. The data were shown to be broadly consistent with the data from the laboratory-based experiment and therefore to some extent the use of a distributed and web-based experiment was validated. Models were developed based on the data and compared with the models derived from the laboratory-based experiment. The effectiveness of additional variables such as age, gender and culture was also investigated thoroughly and modelled. It seems that gender differences do not have considerable influence whilst on the other hand, differences in age and culture are more important. It was noted that the like-dislike bi-polar characteristic behaved differently from the other nine bi-polar characteristics. This is not entirely surprising. Since, like-dislike can probably be categorised under colour preference rather than semiotics. This may suggest that colour preference is different from colour semiotics and may not be modelled in the same way.

8.3 Contributions

The main contributions made by the research in this thesis are as follows:

- A laboratory-based experiment to explore colour semiotics was carried out that broadly supports the earlier work carried out by Ou *et al*.
- A new set of laboratory-based data for colour semiotics have been generated and are available to the community.
- A novel paradigm for carrying out colour semiotic experiments, based on a large-scale Internet presentation, distributed over large numbers of participants was presented, implemented, and to a large extent validated.
- A large set of colour semiotic psychophysical data (over 2000 participants from 58 countries) has been collected and is available to the academic community.

- An analysis of age, gender and cultural differences has been carried out and it has been shown that cultural and age may be significant factors but that gender is probably not.
- An analysis of the so-called inverse problem has been carried out and a solution proposed. The application of this in a colour selector design tool has been explored.

According to the contributions made in this thesis, next sections enlighten the implementations of the results of this research and also possible future work that could be carried out.

8.4 Implementations

So far, some tools have been generated for other properties of colour such as the colour harmony tool which is based on empirical studies and psychophysical experiments [198]. This tool works on the formulated models of colour harmony and shows its applications on cosmetics (see). The conclusions of this study show that the application of colour harmony models may be valid for specific layout designs, since they are based on experiments carried out on single colour patches rather than full images with context. On the other hand, other design elements such as shape, texture and product functionality (which are known to be effective in the colour selection process) are not considered neither in the colour harmony tool either the colour semiotic models. It should be noted that this tool takes an approach towards effective application of the colour semiotics based upon the results of the current research, rather than a deliberate colour selection manual. Otherwise, the developed models and their applications are only generated to make suggestions about colour semiotics and their relationships with colour attributes and personal characteristic. Bear in mind that these models could be updated over time and may change with the collection of more data and variation of age, gender, nationality, etc.

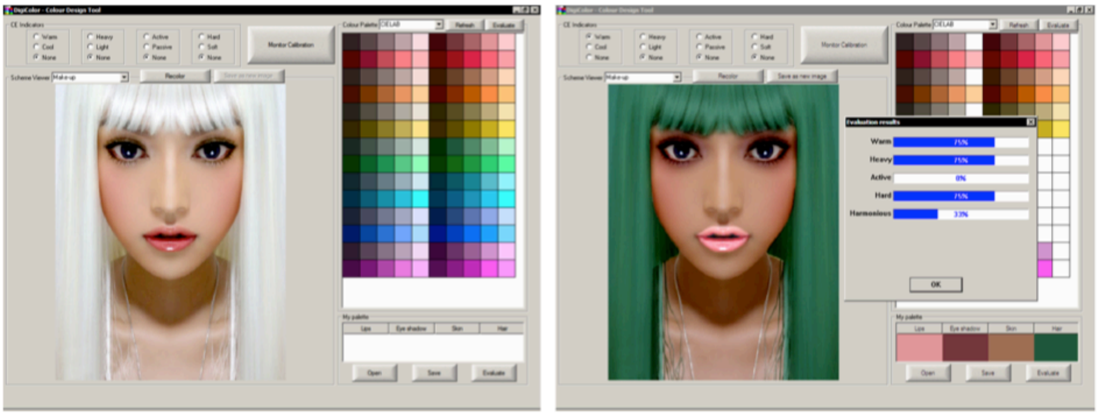


Figure 56: Interface of the colour harmony design tool [198]

8.4.1 Tool layout

Although the tool's main algorithm on which it functions could be very complex and require hard-core programming, the graphical user interface (GUI) should be as simple as possible. A simple approach could be a tool that allows the user to choose a colour from the colour picker and then use the L*a*b* values of the selected colour as inputs of the colour semiotic models (see Figure 57). The outputs of the models are presented by the allocation of the tab on the relevant slider bars for each of the bi-polar characteristics. It may also be useful for the designer to see the selected colour applied to a specific object or shape thus, initially, the colour can be affixed to a certain product by the colour button. This tool is designed to work on the main direction of the model and presents a simple visualisation of how the models work. It allows any user with minimum understanding of colours, semiotics and mathematics to interact with the models easily without having to get involved with them. The status of the slider bars represents the extent to which the chosen colour is perceived for that bi-polar characteristic.

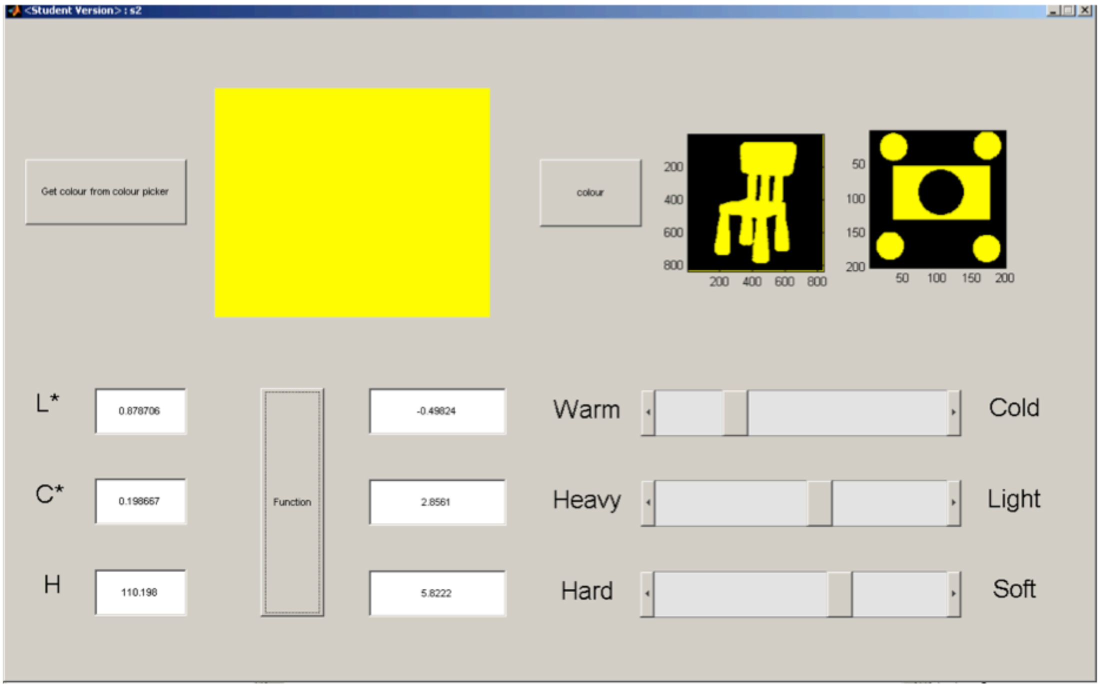


Figure 57: The MATLAB interface of a colour semiotic tool.

It is simple to predict the bi-polar characteristics from a selected colour because that is how the models work. However, in many situations it would be preferable to find a colour (or palette of colours) which best matches a certain set of bi-polar characteristics. Therefore, the inverse-model problem is relevant which was discussed earlier. A primary inverse approach to the models is reflected in a tool illustrated in Figure 40. The user can enter a value for each of the bi-polar characteristics and initially be presented with a relevant range of colours generated by the models. The final interface of the programme is displayed in Figure 35. By this, the user would interact with the tool by moving the slider bars towards the desired positions for each of the characteristics. Corresponding numerical values of the location of the slider bars will be displayed in the box next to it. By hitting the push button, the first four colours resulting from the approximation technique would be displayed in the boxes numbered 1, 2, 3 and 4. These colours may be very similar to each other or may be very different. In addition, the user can choose any of the colours by entering the relevant number in the box indicated at the bottom and the colour would be applied to the set of objects displayed in the images. Therefore, a quick visual presentation of the following colour is displayed on actual objects, which helps the user in identifying the characteristics of that colour on that certain object. Illustration of different objects not only makes a quick connection between the colour and its application, but also makes a comparison between the appearances of colour on different objects. For example a pink shoe can be observed and affixed to feminine and light, while the same colour on a car may not communicate the same characteristics. Of course, in the future Java or flash may be used in order to develop a much more sophisticated interface. Additionally, we must not forget that this is a temporary version of the tool which has been designed to suggest an application of the models derived from the Global Online Survey. In this way the user would apply the formulae to the design process, with more ease and much better understanding, rather than getting confused with the complicated mathematical procedures.

8.4.2 Algorithmic challenges

One of the main challenges of programming the colour semiotic tool is raised by the usage of multiple slider bars. The complexity of this problem is raised regarding to the multi-dimensional nature of the input array of bi-polar characteristics. Considering that the models are developed in isolation from one another and that their interaction together is not identified, it means that the usage of multiple slider bars in the tool can be misinterpreted. The user may think that the generated colour holds all the bi-polar characteristic values. Although certain correlations were identified between the bi-polar characteristics in the previous chapters, it is not possible to conclude a certain descriptive model on how they affect each other. This means that it would not be possible to generate a unique colour by having values for several bi-polar characteristics. This is counted as one of the weaknesses of this tool and seems that the only way to overcome this problem is to take a different approach to the layout of the primary psychophysical experiments. Thus, the combination of bi-polar characteristics should be questioned rather than each in isolation. By that the gathered data would hold information regarding to the combination of bi-polar characteristics and initially may aid the use of multiple slider bars.

Also, the tool requires additional inputs for the age and culture parameters so that the relevant models are called upon selection of each of the variables. The modified interface of the tool is displayed in . As mentioned before, the relevant models are called regarding to the user's selection of the categories. Once the user choses the bi-polar characteristics and allocates the slider bar on the desired value, the inverse of the models is approximated using the approximation technique, resulting in a set of colours that produce the same value for the chosen bi-polar characteristic. Note that as mentioned earlier, the produced colours may not be unique since the approximation technique is based on the iterative process of generating random values for the inverse models. Thus, different colours are generated each time. Also, in cases where the models based on LCH or $LCcos(H° - 50)$ are used; some flipping may occur which is a result of the polar nature of Hue. This flipping would be reflected in the results by the sudden hue shift of the colour sets. Still, the tool offers a flexible interface for the user which will ease the application of the complicated models.

Additionally, in order to test the validity of the models and the colour selector tool, alternative psychophysical experiment would be required [91]. However, it does not fit the framework of this research and would be postponed for further research.

[91] The experiment is run as an interactive MATLAB interface. At the very first stage the participant is displayed with three slider bars and an image of a selected product (watch in the first stage of the experiment then toothbrush and finally sofa). By looking at the black and white image of the product, the participant will have to relocate the bars on the sliders to a value which best presents the characteristics to that product. Also, the anti-characteristics are evaluated. So the opposite of the characteristics that are set by the participants are also evaluated and the colour is derived. The two colours are then applied to the image of the product and displayed to the participant. At this stage, the participant will choose the image with the colour that best represents the characteristics set to the slider bars at the previous stage. If the participant chose the product with the colour derived from the anti-characteristics then that means that the colour semiotic models have not performed well. However, if the participant does choose the image with the colour derived from the colour semiotic models then that means that the models have performed well. This process is repeated for each of the three product image and is also repeated with the same participants after a few days. By repeating this test over time; it will become clear if the participants will choose the same values for the characteristics of the similar product and whether generated colours for different characteristics are accurate.

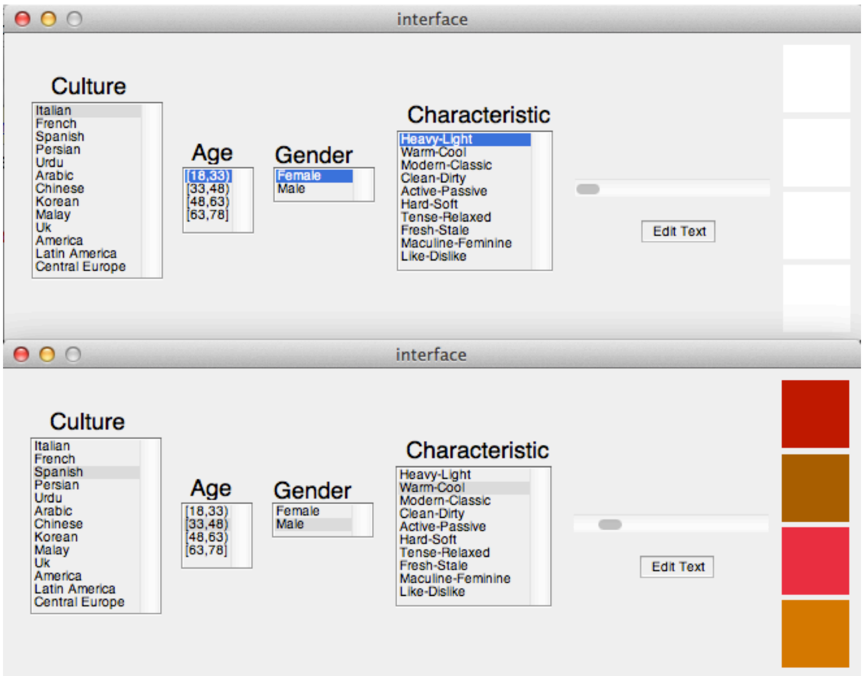


Figure 58: Above: the primary interface of the colour semiotic tool, below: the response to the category selection.

8.5 Future work

In the previous section a software tool is generated with an algorithmic function based upon the models, collecting information from the user as input to the models and generating the outputs of the model in a representative form. The user would apply the models indirectly without getting involved with actual calculations but still addressing the central problem that “without visual aids, verbal colour communication is vague, ill-defined and hardly sufficient for technical, industrial or commercial purposes” [199]. Thus, by creating a tool for colour semiotic models, usage of the colour semiotic models could be more effective. However, the tool suggested in this chapter has a basic interface that is programmed in MATLAB. In the future, a progressed version of this tool could be programmed with a more professional and better looking interface and possibly can be turned into a plug-in for relevant design software such as Photoshop or Illustrator. So that it would be both reachable and useable for all the designers who get involved and challenged with colour selection directly.

References

1. Ou, L.C. A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application*, 2004. **29** pp. 232-240.
2. Moroney, N. Unconstrained web-based color naming experiment. *In Color Imaging VIII* San Jose, 2003. pp. 36-46.
3. Eves, W.R. *The colour Concept generator: A computer tool to propose colour concepts for products*. PhD Thesis: Bournemouth University, 1997.
4. Wyszecki, G. and W.S. Stiles. *Colour science: Concepts and methods, Quantitative data and formulae*, New York: John Wiley & Sons, 1982.
5. Hunt, R.W.G. *Measuring colour*, Chichester, England Fountain Press, 1998.
6. ASTM. 1898 ASTM International Standard Worldwide. West Conshohocken, Pennsylvania.
7. CIE. *In Commission internationale de l'Eclairage proceedings*: Cambridge University Press, 1932.
8. Berns, R.S. *Billmeyer and Saltzman's Principles of Color Technology, 3rd ed*: John Wiley & Sons, 1954.
9. Fairchild, M.D. Psychophysics (Chapter 2), *In Colour appearance models*. New York, John Wiley & Sons, 1997, pp. 150.
10. Fairchild, M.D. *Colour appearance models*, New York: John Wiley & Sons, 1997.
11. Westland, S. and C. Ripamonti. *Computational Colour Science using MATLAB*: John Wiley & sons, 2004.
12. McDonald, R. *Colour physics for industry*. 2nd ed, UK: Textile Institute, 1996.
13. Roberts, D. Instruments and methods for colour measurement, *P. Green and L.W. MacDonald, Editors. In Colour Engineering*, John Wiley & Sons, 2002.
14. Renner, P.B. *Color order and harmony: A color theory for artists and craftsmen*, London: Studio Vista, 1964.
15. CIE. *TC8-01 Technical report, a colour appearance model for colour management systems: CIECAM02*, 2004.
16. Luo, M.R. Development of colour difference formulae. *Review of progress in coloration*, 2002. **32** pp. 28-39.
17. Luo, M.R., G. Cui, and B. Rigg. The development of the CIE2000 colour-difference formula: CIEDE2000. *Colour research and application*, 2001. **26** pp. 340-350.
18. Cui, G., M.R. Luo, B. Rigg, and W. Li. Colour-difference evaluation using CRT colours. Part I: data gathering and testing colour-difference formulae. *Color Research & Application*, 2001. **26** pp. 394-402.
19. Brown, W. and D.L. MacAdam. Visual sensitivity ellipses. *Journal of Optical Society of America*, 1949. **39** pp. 808-834.
20. Davidson, H.R. Ellipses, or Ellipsoids, given in these papers are derived for a stated probability of visually detecting a color difference. *Journal of Optical Society of America*, 1951. **41** pp. 104-111.
21. MacAdam, D.L. Visual sensitivities to color differences in daylight. *Journal of Optical Society of America*, 1942. **32** pp. 247-274.
22. Davidson, H.R. Calculation of Color Differences from Visual Sensitivity Ellipsoids. *Journal of Optical Society of America*, 1951. **41**(12), pp. 1052-1055.
23. CIE *101. Parametric effects in colour-difference evaluation.*, In *CIE Publication* 1993.
24. Hung, P. and R.S. Berns. Determination of constant hue loci for a CRT gamut and their predictions using color appearance spaces. *Color Research & Application*, 1995. **20** pp. 285-295.
25. Morovic, J. and M.R. Luo. Gamut mapping algorithms based on psychophysical experiment. *In 5th IS&T/SID Color Imaging Conference*, Scottsdale, Arizona, 1997. pp. 44-49.

26. MacAdam, D.L. Visual Sensitivities to Color Differences in Daylight. *Journal of Optical Society of America*, 1942. **32**(5), pp.247-273.
27. Brown, W.R.J. and D.L. MacAdam. Visual Sensitivities to Combined Chromaticity and Luminance Differences. 1949. **39**(10), pp.808-834.
28. Fairchild, M.D. Application of colour mixing models to electronic printing. *Journal of electronic imaging*, 1994. **3** pp. 276-287.
29. Johnson, A.J. Methods for characterizing colour printers. *Displays*, 1996. **16** pp. 193-202.
30. Johnson, A.J. Methods for characterizing colour scanners and digital cameras. *Colour Engineering*, 2002.
31. Green, P. *Overview of Characterization methods, in colour engineering*, ed. P. Green and L.W. MacDonald: John Wiley & Sons, 2002.
32. Stokes, M., M. Anderson, S. Chandrasekhar, and R. Motta. 1996. A Standard Default Color Space for the Internet: sRGB.
33. Newman, T. *Improved Color for the World Wide Web: A Case Study in Color Management for Distributed Digital Media*: Printing Industries of America, 2013.
34. CIE. *Parametric effects in colour-difference evaluation*, 1993: CIE publication.
35. Luo, M.R. *MSc programme: colour and imaging science*. Module: colour science for imaging technology, Leeds, UK: University of Leeds, 2011.
36. Bartleson, J. and F. Grin. Optical radiation measurement-visual measurements, Part II: Psychophysical methods. *Academic Express*, 1984. **5**.
37. Agresti, A. *An introduction to categorical data analysis*, Hoboken, NJ: Wiley-Interscience, 2007.
38. Burns, A. and R. Burns. *Basic marketing research* 2nd ed, New Jersey: Pearson Education, 2008.
39. Dawes, J. Do data characteristics change according to the number of scale points used? an experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research*, 2008. **50**(1) pp. 61-77.
40. Trochim, W.M. *Likert scaling*. Research methods knowledge base: Atomic Dog Publishing, 2001.
41. Torgerson, W.S. *Theory and methods of scaling*, Michigan: Wiley, 1958.
42. Thurstone, L.L. A law of comparative judgment. *Psychology Review*, 1927. **34** pp. 273-286.
43. Snider, J.G. and C.E. Osgood *Semantic differential technique : A Sourcebook*, Chicago: Aldine Pub. Co, 1969.
44. Thompson, B. *Exploratory and confirmatory factor analysis: understanding concepts and applications.*, Washington, DC: American Psychological Association, 2004.
45. Costello, A.B. and J. Osborne. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 2005. **10**(7) pp. 1-9.
46. Hair, J., R. Anderson, R. Tatham and W.C. Black. *Multivariate data analysis.* 4th ed, New Jersey: Prentice-Hall, 1995.
47. Frees, E.W. *Data analysis using regression models : the business perspective / Edward W. Frees*, Upper Saddle River, New Jersey: Prentice Hall, 1996.
48. Siegel, S. *Nonparametric statistics for the behavioural sciences*, New York: McGraw-Hill, 1988.
49. Milton, H. *Packaging design*, London: The Design Council, 1991.
50. Ou, L. and M.R. Luo. A colour-harmony model for two-colour combinations. *Color research and application*, 2006. **31**(3) pp. 191-204.

51. Sueeprasan, S., P. Srimork, A. Hansuebsai, T. Sato, and A. Pungrassamee. Quantitative analysis of Thai sensation on colour combination. *In 10th Congress of the International Color Association* Granada, Spain, 2004. pp. 271-274.
52. Guildford, J. and P. Smith. A system of color preferences. *American Journal of Psychology*, 1959. **72**(4) pp. 487-502.
53. Palmer, S. and K. Schloss. An ecological valence theory of human color preference. *Proceedings of the National Academy of Science*, 2010. **107**(19) pp. 8877-8882.
54. Cardello, A.V., H. Schutz, C. Snow, and L. Lesher. Predictors of Food Acceptance, Consumption and Satisfaction in Specific Eating Situations. *Food quality and preference*, 2000. **11** pp. 201-216.
55. Kennedy, O., B. Stewart-Knox, P. Mitchell, and D. Thurnham. Flesh colour dominates consumer preference for chicken. *Appetite*, 2005. **44** pp. 181-186.
56. Khliji, S., R. van de Ven, T. Lamb, M. Lanza, and D. Hopkins. Relationship between consumer ranking of lamb colour and objective measures of colour. *Meat Science*, 2010. **85** pp. 224-229.
57. Williams, S., T. Frye, M. Frigg, D. Schaefer, K. Scheller, and Q. Liu. Vitamin E. *Meat Focus International*, 1992. **3**(2) pp. 22-23.
58. Ambrose, G. and P. Harris. *Packaging the Brand- The Relationship between Packaging Design and Brand Identity* Lausanne: AVA Publishing SA, 2011.
59. Lindstorm, M. *Brand Sense: Sensory Secrets Behind the Stuff We Buy*, London: Kogan Page, 2005.
60. Westland, S., K. Laycock, V. Cheung, P. Henry, and F. Mahyar. Colour harmony. *Colour design and creativity*, 2007. **1**(1) pp. 1-15.
61. Holtzschue, L. *Understanding color*: John Wiley & Sons, 2006.
62. Harkness, N. The colour wheels of art, perception, science and physiology. *Optical Laser Technology*, 2006. **38** pp. 219-229.
63. Mahyar, F. *Colour harmony in the context of design*,PhD Thesis In *School of Design*: University of Leeds, 2010.
64. Wong, W. *Principles of color design*: John Wiley & Sons, 1997.
65. Kuehni, R. *An introduction to practice and principles of color*: John Wiley & Sons, 2005.
66. Pressey, S.L. The influence of colour upon mental and motor efficiency. *American Journal of Psychology*, 1921. **1**(32) pp. 326-356.
67. Dubrovner, R.J., W.J. Lackum, and H. Jost. A study of the effect of color on productivity and reaction time in the rorschach test. *Journal of clinical psychology*, 1950. **6** pp. 331-336.
68. Goodfellow, R.A.H. and P.C. Smith. Effects of environmental color on two psychomotor tasks. *Percept Mot Skills*, 1973. **37** pp. 296-298.
69. Maga, J. Influence of colour on taste thresholds. *Chemical Senses & Flavour*, 1974. **1** pp. 115-119.
70. Guéguen, N. The effect of glass colour on the evaluation of a beverage's thirst-quenching quality. *Current Psychology Letters*, 2003. **11**(2) pp. 2-6.
71. Rubinoff, M. and D. Marsh Candidates and colors: an investigation. *Perceptual and Motor Skills*, 1980. **50** pp. 868-870.
72. Garrett, J. and C. Brooks. Effect of ballot colour, sex of candidate, and sex of college students of voting age on their voting behaviour. *Psychological Reports*, 1987. **60** pp. 39-43.
73. Read, M., A. Sugawara, and J. Brandt. Impact of space and color in the physical environment on preschool childrens' cooperative behaviour. *Environment and Behavior*, 1999. **31** pp. 413-428.

74. Spath, D. The psychological and physiological effect of ‘Cool Down Pink’ on human behavior. *In AIC International Conference*, Zurich, Switzerland, 2011. pp. 751-754.
75. Wise, B. and J. Wise*. The human factors of color in environmental design: a critical review*, In *Final NASA Contractor Report No. NCC 2-404* 1988: NASA.
76. Mikellides, B. Color and physiological arousal*. Architectural and planning research*, 1990. **7**(1) pp. 13-19.
77. Walters, J. and M.J. Apter Color preference arousal, and the theory of psychological reversals*. Motivation and emotion*, 1982. **6**(3) pp. 193-215.
78. Berlyne, D.E. *Conflict, arousal and curiosity*, In *McGraw-Hill* 1960. New York.
79. Apter, M.J. Some data inconsistent with the optimal arousal theory of motivation*. Perception and Motivation Skills*, 1976. **43** pp. 1209-1210.
80. Apter, M.J. On the concept of bistability*. International Journal of General Systems*, 1981. **6** pp. 225-232.
81. Apter, M.J. *The experience of motivation: the theory of psychological reversals*, In *Academic Press* 1982. New York.
82. Eysenck, H.J. and M.W. Eysenck*. Personality and individual differences: A natural science approach*, 1985. Plenum, New York.
83. Gray, J.A. *The Neuropsychologyo f Anxiety: An Inquiry into the Functions of the Septo-Hippocampal sytem*, In *Oxford University Press* 1982. Oxford.
84. Henthrone, T.L., M.S. Latour, and R. Nataraajan. Fear appeals in print advertising: an analysis of arousal and ad response*. Journal of Advertising*, 1993. **22**(2) pp. 59-69.
85. Smith, K.C.P. and M.J. Apter. *A theory of psychological reversal*, 1975. Picton, Chippenham, England.
86. Thayer, R.E. Toward a psychological theory of multidimensional activation *Motivation and emotion*, 1978. **2**(1) pp. 1-33.
87. Gorn, G.J., A. Chattopadhyay, T. Yi, and D.W. Dahl. Effects of color as an executional cue in advertising: they're in the shade*. Management Science*, 1997. **43**(10) pp. 1387-1997.
88. Kwallek, N., K. Soon, and C.M. Lewis. Work Week Productivity, Visual Complexity, and Individual Environmental Sensitivity in Three Offices of Different Colour Interiors*. Color Research and Application*, 2006. pp. 14.
89. Wright, A. *The beginner's guide to colour psychology*, London, UK: Colour Affects, 1998.
90. Adams, F.M. and C.E. Osgood. A cross - cultural study of the affective meanings of color*. Journal of Cross - Cultural Psychology*, 1973. **4** pp. 135-156.
91. Valdez, P. and J. Mehrabian. Effect of color on emotions. *Journal of Experimental Psychology: general*, 1994. **123**(4) pp. 394-409.
92. Mehrabian, A. *Manual for the questionnaire measure of stimulus screening and arousibility*, 1976. Los Angeles, CA: UCLA.
93. Mehrabian, A. Theory and evidence bearing on a scale of trait arousability *Current Psychology Letters*, 1995. **14** pp. 3-28.
94. Mehrabian, A. and J.A. Russell*. An approach to Environmental Psychology*, 1974. Cambridge, Massachusetts: MIT press.
95. Kruithof, A.A. Tubular luminescence lamps for general illumination*. Philips Technological Review*, 1941. **6**(3) pp. 65-73.
96. Mukae, H. and M. Sato. The effect of colour temperature of lighting sources on the autonomic nervous functions*. The Annals of Physiological Anthropology*, 1992. **11**(5) pp. 533-538.

97. Lee, T.R. and V. Sun. Investigating the psychological and physiological effects of human exposure to color light. *In AIC International*, Zurich, Switzerland, 2011. pp. 526-529.
98. Krippendorf, K. On the essential contexts of artefacts or on the proposition that "design is making sense (of things)" *Design issues*, 1989. **5**(2) pp. 9-39.
99. Bijker, W.E. and J. Law. *Shaping Technology*, Massachusetts: Massachusetts Institute of Technology, 1992.
100. Pearman, H. *Urban Approaches* 1993: Blueprint.
101. Ulrich, K.T. and S.D. Eppinger. *Product design and development*, New York McGraw-Hill Higher Education 2008.
102. Shiba, S. *Four practical revolutions in management: systems for creating unique organizational capability Portland*, Cambridge: Productivity Press, 2001.
103. Baxter, M. *Product design: a practical guide to systematic methods of new product development* London Chapman & Hall, 1995.
104. Schmitt, B.H. and Y. Pan. Managing corporate and brand identities in the asian-pacific region. *California management review*, 1994. **36** pp. 32-48.
105. Holden, T.J.M. The color of meaning: The significance of black and white in television commercials. *Interdisciplinary information sciences*, 1997. **3**(2) pp. 125-146.
106. Glaser, M. *I listen to the market*. Marshall Blonsky, ed. On Signs, Oxford: Blackwell, 1985.
107. Bohle, R. and M. Garcia. *Readers reactions to color in newspapers*, In *Annual Meeting of the Association for Education in Journalism and Mass Communication* 1986.
108. Brandt, E.R. The memory value of advertisements. *Archives of Psychology*, 1925. **8**(79).
109. Dooley, R.P. and L.E. Harkins. Functional and attention-getting effects of colour on graphic communications. *Perceptual and Motor Skills*, 1970. **31** pp. 851-854.
110. Shanon, C. and W. Weaver. *The mathematical theory of communication*, In *University of Illinois Press* 1949.
111. Chandler, D. *Semiotics: the basics*, London: Routledge, 2002.
112. Peirce, C.S. *Collected writings*, In *MA: Harvard University Press.* 1931. Cambridge.
113. Kenney, K. *Representation Theory*. Smith, K. Handbook of Visual Communications: Theory, Methods and Media, ed. L.E. Associates New Jersey: London 2005.
114. Hyman, J. Pictorial art and visual experience. *British Journal of Aesthetics*, 2000. **40** pp. 21-45.
115. Eco, U. *How culture conditions the colours we see*. The Communication Theory Reader Paul Cobley ed. Routledge London, 1996.
116. Eco, U. *Towards a semiotic enquiry into the television message*. 1980, ed. Corner. and Hawthorn., 1965.
117. Hupka, R.B., Z. Zaleski, J. Otto, L. Reidl, and N.V. Tarbarina. The colors of anger, envy, fear, and jealousy. A cross - cultural study. *Journal of Cross - Cultural Psychology*, 1997. **28** pp. 151-171.
118. Kargere, A. *Color and personality*, New York, NY: Philosophical Library, 1949.
119. Biren, F. Colour preference as a clue to personality. *Art Psychotherapy*, 1973. **1** pp. 13-16.
120. Heath, R.P. *The wonderful world of color; Marketing tools*. 1997 [cited 2013 October]; Available from: http://www.demographics.com/publications/mt/97_mt/9710_mt?mt971022.tm.

121. Lane, R. Does orange mean cheap? *Forbes Journal*, 1991. **148**(December) pp. 144-147.
122. Manav, B. Color-emotion associations and color preferences. *Color research and application*, 2006. pp. 8.
123. Taft, C. Color meaning and context: Comparisons of semantic ratings of colors on samples and objects. *Color research and application*, 1997. **22**(1) pp. 40-50.
124. Xin, J.H., K.M. Cheng, G. Taylor, T. Sato, and A. Hansuebsai. Cross - regional comparison of colour emotions, Part 2: Qualitative analysis. *Color research and application*, 2004. **29** pp. 8.
125. Sato, T., K. Kajiwara, J.H. Xin, A. Hansuebsai, and J. Nobbs. Numerical expression of colour emotion and its application. *In International Color Association, Proceedings of AIC Color' 2001*, Rochester, New York, . 2001. pp. 409-412.
126. Osgood, C.E. The cross - cultural generality of visual - verbal synaesthetic tendencies. *Behavioural Sciences* 1960. **5** pp. 23.
127. Arnold, M.B. *Emotion and personality*, 1960. New York: Columbia University Press.
128. Izard, C.E. *Patterns of emotions; A new analysis of anxiety and depression* 1972. New York: Academic Press.
129. Plutchitt, R. *Emotion: A psycho-evolutionary synthesis*, 1980. New York: Harper and Row.
130. Mazois, F. Reconstruction of a wall,'Edifice' of Eumachia, Pomeii, from Les Ruines De Pompeii. 1929.
131. Solli, M. and R. Lenz. Color emotions for image classification and retreival. *In 4th European Conference on Colour in Graphics, Imaging, and Vision, CGIV*, . 2008.
132. Tarasewich, P. and K.S. Nair. Designer-moderated product design. *IEEE Transactions on Engineering Management*, May 2001. **48**(2).
133. Burnswick, E. *The conceptual framework of psychology*, 1952. Chicago: University of Chicago Press.
134. Geistfeld, L.V., G.B. Sproles, and S.B. Badenhop. The concept and measurement of a hierarchy of product characteristics. *Advances in Consumer Research*, 1977. **4**(1) pp. 302-307.
135. Kaul, A. and V.R. Rao. Research for product positioning and design decisions: an integrative review. *International journal of research in Marketing*, 1994. **12**(4) pp. 293-320.
136. Snelders, D.H. *Subjectivity in the consumers judgment of products*. PhD Thesis, Delft: Delft University of Technology, 1995.
137. Veryzer, R.W. A nonconscious processing explanation of consumer response to product design. *Psychology & Marketing*, 1999. **16**(6) pp. 497-522.
138. Kirmani, A. Advertising repetition as a signal of quality: If it's advertised so much, something must be wrong. *Journal of advertising*, 1997. **26**(3) pp. 77-86.
139. McCracken, G. *Culture and consumption*, In *Bloomington IN: Indian University Press* 1988.
140. Chang, W.C. and T.Y. Wu. Exploring types and characteristics of product forms. *International Journal of Design*, 2007.
141. Hertenstein, J.H., M.B. Platt, and R.W. Veryzer. The impact of design effectiveness on corporate financial performance. *Journal of Product Innovation Management*, 2005. **22**(1) pp. 3-21.
142. Lewalski, Z.M. Product aesthetics: An interpretation for designers. *Design & Development Engineering Press*, 1988.

143. Yamamoto, M. and D.R. Lambert. The impact of product aesthetics on the evaluation of industrial products. *Journal of Product Innovation Management*, 1994. **11**(4) pp. 309-324.
144. Bloch, P.H. Seeking the ideal form-Product design and consumer response. *Journal of Marketing Practice*, 1995. **59**(3) pp. 16-29.
145. Blijlevens, J., E.H. Creusen, and J.P.L. Schoormans. How consumers perceive product appearance: The identification of three product appearace attributes. *International Journal of Design*, 2009. **3**(3) pp. 27-35.
146. Diane, T. and T. Cassidy. *Colour Forecasting* Blackwell Publishing, 2005.
147. Chi, M.T.H., P.J. Feltovich, and R. Glaser. Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 1981. **5**(2) pp. 121-152.
148. Tanaka, J.W. and M. Taylor. Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 1991. **23**(3) pp. 457-482.
149. Hsu, S.H., M.C. Chuang, and C.C. Chang. A semantic differential study of designers' and users' product form perception. *International Journal of Industrial Ergonomics*, 2000. **25**(4) pp. 375-381.
150. Russell, D. *Colour in industrial design*, London: Design Council, 1991.
151. Ellis, S.R. *A psychometric investigation of a scale for the evaluation of the aesthetic element in consumer durable goods*, In *University of Arizona* 1993. Tucson, AZ: University of Arizona.
152. Henderson, P.S., J.L. Giese, and J.A. Cote. Impression management using typeface design. *Journal of Marketing Practice*, 2004. **68**(4) pp. 60-72.
153. Orth, U.R. and K. Malkewitz Holistic. Package design and consumer brand impressions. *Journal of Marketing*, 2008. **72**(3) pp. 64-81.
154. Alsop, R. and C.E. Osgood. Color grows more important in catching consumers' eyes. *The Wall Street Journal*, 1984. **29** pp. 37.
155. Achauss, A. Tranquilizing effect of color reduces aggressive behavior and potential violence. *Journal of Orthomolecular Psychiatry*, 1979. **8** pp. 218-221.
156. Grimes, A. and I. Doole Exploring the relationships between colour and international branding: A cross cultural comparison of the UK and Taiwan. *Journal of Marketing Management*, 2010. **14**(7) pp. 799-817.
157. Tucker, J.V. Printing: psychology of color. *Target Marketing*, 1987. **10**(7) pp. 40-49.
158. Geboy, L.D. Colour makes a better message. *Journal of Health Care Marketing*, 1996. **16**(summer) pp. 52-54.
159. Mortan, J. *Colors that sell, Tried and tested color schemes*. 2013; Available from: http://www.colorvoodoo.com/cvoodoo9_colorsthatsell.html.
160. Sato, T., K. Kajiwara, H. Hoshino, and T. Nakamura. Quantitative evaluation and categorising of human emotion induced by colour. *Advertisement and Colour Science Technology*, 2000. **3** pp. 53-59.
161. Xin, J.H. and K. Cheng. Quantitative evaluation of colour emotion. *JSPS Fund Meeting.*, 2000. pp. 71-86.
162. Osgood, C.E., G.J. Suci, and P.H. Tannenbaum. *The measurement of meaning*, In *University of Illinois Press* 1957.
163. Spearman, C. General intelligence: Objectively determined and measured. *American Journal of Psychology*, 1904. **15** pp. 201-293.
164. Wright, B. and L. Rainwater. The meannings of colour. *J Gen Psychol*, 1962. **67** pp. 89-99.
165. Hogg, J. A principle component analysis of semantic differential judgements of single colours and colour pairs. *J Gen Psychol*, 1969. **80** pp. 129-140.

166. Hogg, J., S. Goodman, T. Porter, B. Mikellides, and D. Preddy. Dimensions and determinants of judgments of colour samples and a simulated interior space by architects. *Joournal of Psychology*, 1979. **70** pp. 231-242.
167. Kobayashi, S. The aim and method of the color image scale. *Color Research & Application*, 1981. **6** pp. 93-107.
168. Crozier, W. The psychology of colour preference. *Rev pro coloration*, 1996. **26** pp. 63-72.
169. Ou, L.C., M.R. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. Part 1: Colour emotions for single colours. *Color research and application*, 2003. **29**(3) pp. 232-240.
170. Richerson, P.J. *Not by genes alone: how culture transformed human evolution* Chicago, USA: Library of Congress Cataloging in Publishing Data, 2005.
171. Mohammadazadeh, M. and S. Westland [cited 2012; Available from: http://www.keysurvey.co.za/.
172. Scott, W.A. Reliability of content analysis: The case of nominal scale coding. *Public quarterly*, 1955. **19** pp. 321-325.
173. Cohen, J. A coefficient of agreement for nominal scales. *Educational and psychological measurement.*, 1960. **20** pp. 37-46.
174. Fleiss, J.L. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 1971. **76**(378-382).
175. Cohen, J. Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 1968. **70** pp. 213-220.
176. Light, R.J. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol. Bull*, 1971. **76** pp. 365-377.
177. Bishop, D.E. *Probability and statistics: models for research* New York: John Wiley, 1971.
178. Tabachnick, B.G. and L.S. Fidell. *Using multivariate statistics*, New York: Harper Collins, 1996.
179. Brace, N. *SPSS for psychologists*. 5th ed, Basingstoke: Palgrave Macmillan, 2012.
180. Kaiser, H. A second generation Little Jiffy. *Pshchometrika*, 1970. **35** pp. 401-415.
181. Kaiser, H. An index of factorial simplicity. *Pshchometrika*, 1974. **39** pp. 31-36.
182. Bartlett, M.S. A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society*, 1954. **16**(Series B) pp. 296-298.
183. Cattel, R.B. The scree test for number of factors. *Multivariate behavioural research*, 1966. **1** pp. 245-276.
184. Westland, S. *Personal communication "Approximation techniques"*, M. Mohammadazadeh, Editor 2012. Leeds.
185. Morovic, J. *Color gamut mapping*, Chichester: John Wiley & Sons, 2008.
186. MacAlpine, P. *Colour analysis*. [cited 2012; Available from: http://www.clear.rice.edu/elec301/Projects02/artSpy/color.html.
187. Beretta, G. and N. Moroney. Validating large-scale lexical color resources. *In AIC Conference*, Zurich, Switzerland, 2011.
188. Zuffi, S., P. Scala, C. Brambilla, and G. Beretta. Web-based vs. controlled environment psychophysics experiment. *In Image Quality and System Performance IV*, 2007. pp. 49407.
189. Mylonas, D., L. MacDonald, and S. Wuerger. Towards an Online Color Naming Model. *In Eighteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*, San Antonio, Texas, 2010. pp. 140-144.

190. Munroe, R. and E. Ellis [cited 2011 March 11]; Available from: http://blag.xkcd.com/.
191. *Internet world stats*. [cited 2013; Available from: http://www.internetworldstats.com/stats.htm.
192. Mohammadazadeh, M. *Colour semiotics*. [cited 2013 January]; Available from: http://coloursemiotics.wordpress.com/.
193. Beretta, G. [cited 2011 June]; Available from: http://www.mostlycolor.ch/2011/06/color-semiotics.html?spref=fb.
194. Parraman, C. [cited 2011 June]; Available from: http://www.uwe.ac.uk/sca/research/create/opportunities/index.html.
195. *Survey galaxy*. [cited 2011; Available from: http://www.surveygalaxy.com/info.asp?action=gPromoting.
196. Mohammadazadeh, M. http://prezi.com/xycp6ksxu-lx/colour-semiotics/. [cited 2013 January].
197. Morgia, M. [cited 2012; Available from: http://monicadmurgia.com/2011/06/23/color-semiotics-unraveling-hidden-meaning-of-color/.
198. Ou, L.C. and M.R. Luo. A Colour Design Tool Based on Empirical Studies. *In Proceedings of the Design Research Society Conference* Sheffield, UK, 2008.
199. Hirschler, R. Colour communication in industry from design to product - with special emphasis on textiles, *C. Parraman, Editor In Colour-coded*. Bristol, Society of Dyers and Colourists, 2011, pp. 263-273.

Appendix

The algorithm of the categorical-judgment which is programmed in MATLAB is displayed in the following.

Frequency Matrix with 5 categories

	Categories				
Colours	1	2	3	4	5
A	2	0	1	1	0
B	2	0	2	0	0
C	1	0	1	1	1

$$CF_m = \sum_{m=1}^{m} F_m$$

Cumulative Frequency (CF)

	Categories				
Colours	1	2	3	4	5
A	1	1	2	3	4
B	2	2	3	4	4
C	2	2	4	4	4

$$CP = CF/N$$

Cumulative Proportion Matrix (CP)

	Categories				
Colours	1	2	3	4	5
A	0.25	0.25	0.50	0.75	1
B	0.50	0.50	0.75	1	1
C	0.50	0.50	1	1	1

$$LG = \ln\frac{(CP * N + 0.5)}{(N - CP * N + 0.5)}$$

Logistic Function Matrix (LG)

	Categories				
Colours	1	2	3	4	5
A	-0.85	-0.85	0	0.85	2.20
B	0	0	0.85	2.20	2.20
C	0	0	2.2	2.20	2.20

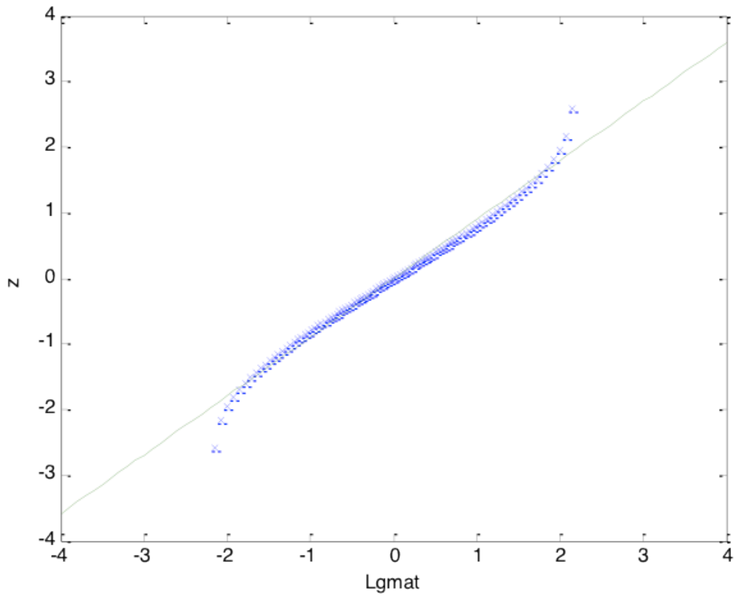


Figure 59: Linear relationship between z values and LG values which are simulated with N=4 for this example is 0.897.

$Z = LG * P$

Z-score Matrix (Z)

	Categories				
Colours	1	2	3	4	5
A	-0.76	-0.76	0	0.76	1.97
B	0	0	0.76	1.97	1.97
C	0	0	1.97	1.97	1.97

$DM_m = Z_{m+1} - Z_m$

Difference Matrix (DM)

	Categories				
Colours	1	2	3	4	
A	0	0.76	0.76	1.21	
B	0	0.76	1.21	0	
C	0	1.97	0	0	
Mean 1	0	1.164	0.65	0.4	
Boundaries	0	0	1.16	1.82	2.[illegible]3
Category	1	2	3	4	5

Mean of each column

$Boundary_1 = 0$

$Boundary_m = Boundary_{m-1} + Mean_{m-1}$

Scale Matrix (SM)

$S = Boundary - Z$

	Categories						
Colours	1	2	3	4	5	Mean 2	Category
A	0.76	0.76	1.164	1.06	0.25	0.80	3
B	0	0	0.40	-0.15	0.25	0.10	3
C	0	0	-0.81	-0.15	0.25	-0.14	1

$Boundary_2 < Mean\ 2 < Boundary_4$