Investigation into the relationship between visual perception and instrumental colour measurement of teeth

Author: Chelsea Rose Sullivan

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Design

January 2022

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Research from this thesis has been published in the Journal of Dentistry called “A Yellowness Index for use in Dentistry” in collaboration with Colgate-Palmolive. Perceptual response data from 500 participants collected by Dr. QianQian Pan, Professor Stephen Westland and Roger Ellwood was available for use in the development of the yellowness index. Their data was used in this thesis as well as in the published paper. The index and validation data was developed by the author.

Dr. QianQian Pan, Professor Stephen Westland and Roger Ellwood collected perceptual yellowness responses for a selection of 52 samples in chapter 5. This data was not published and was available for use in this thesis. This data was compared to instrumental colour measurements collected by the author. All statistical analysis was done by the author.

All data collected by Dr. QianQian Pan, Professor Stephen Westland and Roger Ellwood was for work in collaboration with Colgate-Palmolive.

Acknowledgements

This project would not have been possible without the financial support and exceptional resources from Colgate-Palmolive and the School of Design at the University of Leeds. Thank you to this PhD program for providing new experiences, friends, and countless challenges.

Professor Stephen Westland, I will forever be grateful for your constant encouragement, patience, support and knowledge throughout this project. More importantly, thank you for believing in me and taking a risk by providing this opportunity to grow during a challenging time in my personal life. Thank you for helping me to build more confidence as a researcher.

Special thanks to Roger Ellwood and Dr. Richard Hogan from the Colgate-Palmolive Research and Development team for your kindness and knowledge. In addition, thank you to the Dental Health Unit and Colgate for allowing me to use your research facilities. Thank you to my second supervisor Dr. Kiada Xiao for providing me additional experiences and opportunities during my PhD.

Endless appreciation and love to my entire family for always encouraging and supporting wherever my life takes me. The last six years have been mentally, physically and emotionally challenging, but your love has helped me continue to grow, adventure and be fearless. Mom, Dad and Momma Sue, thank you for always making sure I am taken care of and live comfortably. Thank you to my brothers, Nick and Tyler, my sister-in-law, Jill for inspiring me and being great roll-models. You all have helped me accomplish many things in my life including this PhD.

Thank you to Joe, for being by my side through the many challenges I have faced, making me smile, and always being patient with me, my mood swings, and health issues. In addition, thanks for listening me complain about how much I hate writing. Special shout out to my dog, Teddy, for giving me all the love, cuddles, comfort and being the ultimate exercise buddy during this chapter of our lives.

Thank you to my beautiful best friends for bringing constant smiles, coffee dates, walks, laughter, lasting memories, listening ears and/or shoulders to cry on. I'm beyond fortunate to have you in my life: Becca, Alison, Lissie, Jie, Abby, Dylan, Mel, Mark, Alex N., Ryan, Arthur, Emily, Katie, Nani, and Bintan.

Thank you to other friends who have helped me through different stages of life: Evan, Mirachelle, Devin, My Bark Park family, Arran, Alex W., Helena, and Louise.

Algy Kazlauciunas, thank you for encouraging me to 'colour outside the lines' during my MSc.

Thank you to VeriVide for supporting me while I finished this PhD.

Abstract

Individuals desire to seek aesthetic dental treatments such as restoration, dental fillings, and whitening treatments to obtain the 'perfect' smile has given rise to research to determine the most accurate method for tooth colour measurement. Accurate tooth colour measurement is essential in producing satisfactory results for restoration and fillings, improving communication between dentist and patient, as well as determining the efficacy of tooth whitening products. Currently there is no 'gold' standard method for assessing tooth colour. Gloss, curvature, small size, in vivo location and translucency are attributes of teeth which make accurate colour measurement difficult. The aim of this thesis is to evaluate 'accuracy' of tooth colour measurement by exploring some of the issues arising in visual perception, indices, and instrumental colour measurement. A yellowness index (YIO) is developed to assess changes in perceptual yellowness instead of whiteness which highlighted issues in indices and perceptual evaluation. The amount of edge-loss that occurs in the spectrophotometer, spectroradiometer, Vita EasyShade and digital camera was assessed to determine the amount of error produced by each measurement device. A novel approach to the assessment of accuracy is introduced whereby accuracy is expressed by the instrument's ability to 'agree' with visual perception, with perceptual responses as the 'reference'. Statistical analysis methods r^2, percent wrong decisions and STRESS analysis are used in order to assess the correlation between instrumental measurement and visual perception responses (Z-scores). It was found that there is no 'gold' standard instrument, but there are indices that produce better correlation between instrumental measurement and visual perception. WIO and YIO produced the best results for the assessment in changes in tooth whiteness and tooth yellowness no matter which instrument was used. The Vita Easy Shade was found to perform poorly in comparison to all other instruments.

Table of Contents

Table of Contents

List of Tables

List of Figures

1 Literature Review

1.1 Introduction

Smile! A simple facial expression where the corners of an individual's mouth are turned up and teeth are exposed. It denotes self-esteem, self-confidence, an individual's oral health, social class, age and contributes to one's overall attractiveness (Joiner and Luo, 2017; Montero *et al.*, 2014; Höfel *et al.*, 2007; Xiao *et al.,* 2007). People are generally perceived more favourably in regards to happiness, social relations, relationship status and academic performance with healthy dentition as opposed to heavily stained dentition (Montero *et al.*, 2014; Kershaw *et al.*, 2008). Desire for a white smile has been important since the middle ages, yet today's media has globally influenced consumers perception of what is socially deemed as the perfect white 'Hollywood' smile (Schmidseder, 2000). The idealised standards of dental appearance can be internalised by consumers, raising self-awareness of one's own tooth colour which can lead to dissatisfaction with one's own appearance (Kovacevic Pavicic *et al.*, 2017; Carey, 2014; Alkhatib *et al.*, 2005; Kershaw *et al.*, 2008; Höfel *et al.*, 2007). Dissatisfaction of tooth colour appearance has been shown to affect individuals world-wide with a rate of 34% of 180 adult participants in the United States, 52.6% of 405 participants in China, 65.9% of 220 students in Saudi Arabia, 56.2% of 235 participants in Malaysia (Odioso *et al.*, 2000; Xiao *et al.*, 2007; Al-zarea, 2013; Tin-Oo *et al.*, 2011, Montero *et al.,* 2014; Alkhatib *et al.*, 2005). Dissatisfaction with tooth colour appearance has been shown to be associated with increased desire to improve dental aesthetics, specifically 'whitening' treatments (Joiner and Luo, 2017; Al-Zarea, 2013). However, tooth 'whitening' is not the only aesthetic treatment available to reach a perfect smile. Aesthetic dental treatments include tooth colour fillings, crowns, restorations, bleaching, dentures, veneers, and orthodontic treatments (Al-Zarea, 2013). Treatments that improve aesthetics have been found to increase a patient's quality of life,

social integration, and psychological status (Tin-Oo *et al.*, 2011; Joiner, 2004; Kovacevic Pavicic *et al.*, 2017). The desire to increase one's dental standards has inadvertently decreased the severity of dental caries emerging in areas with healthy populations in developed countries (Montero *et al.*, 2014).

The goal of dentists is to aid patients to reach an accepted level of satisfaction with their dentition regarding oral health and aesthetics (Al-Zarea, 2013; Joiner, 2004; Vallittu *et al.*, 1996). Communication of expectations between patients and dentists as well as dentists and laboratories is critical in order to prevent unrealistic promises (Höfel *et al.*, 2007; Kovacevic Pavicic *et al.*, 2017). Tooth colour specification aids in evaluating the efficacy of bleaching treatments (at home and in-office), choice of filling colour, veneers and selection of appropriate materials in desired shade for restorations (Oguro *et al.*, 2016; Johnston, 2009; Curd *et al.*, 2006). For restorations, wrong colour assessment can result in restoration instalment removal for colour correction leading to potential for oral structure damage as well as additional expenses due to additional required time for patients, dentists and dental technicians (Brandt *et al.,* 2017). However, accurately assessing tooth colour is difficult.

Interest in colour measurement research in dentistry has significantly increased since 1970. The number of published papers citing 'colour' and 'dentistry' increased from 107 (1970) to 5400 (2014) (Chu *et al.*, 2010; Khashayar *et al.,* 2014). This research area continues to increase with the number of papers rising to 12275 in 2021 in areas such as prosthodontics, aesthetics and dental materials science, and perceptible/acceptable colour quantification (Khashayar *et al.*, 2014).

Figure 1: The relationship between visual perception and instrumental measurement

The aim of colour measurement research in dentistry is to express visual results in an quantitative way, which is done using indices as shown in Figure 1 (ASTM E313-05, 2005). Visual assessment and colour measurement devices such as spectrophotometers, colorimeters, spectroradiometers and digital cameras, have been used to measure tooth colour. Visual assessment is subjective; it can vary depending on observer and viewing conditions, and can be constrained by the shade guide tabs that are typically used as references (Tin-Oo *et al.*, 2011; McLaren, 1970; Khurana *et al.*, 2007). While colour measurement devices produce a more objective method for tooth colour assessment, these instruments do produce their own set of problems such as reproducibility of the measurement itself due to the irregular surface of teeth, edge-loss, size of teeth, being multilayered, in-vivo location, and exhibit colour transitions in all directions (gingival to incisal, messiah to distal and labial to lingual (Johnston, 2009; Chu *et al.,* 2010). Numerous devices are available, but there is no universal 'gold' standardised method for tooth colour measurement. The standard that was published by ASTM (ASTM E2466) in 2013 using a digital camera was subsequently withdrawn. Even prior to this standard being retracted, research was still being investigated to deduce the most accurate method for measuring tooth colour. Numerous indices are available for evaluating tooth colour; it is not clear which one best relates instrumental measurement to visual perception.

The aim of this thesis is to evaluate 'accuracy' of tooth colour measurement by assessing some of the issues arising in visual perception, indices, and instrument measurement. The first experiment

describes the development of a ‘yellowness’ index to assess changes in perceptual ‘yellowness’ instead of 'whiteness’ which evaluates issues that arise in indices as well as perceptual evaluation. The second experiment investigates the amount of edge-loss that occurs in a range of colour measurement instruments available to assess the amount of error produced by each measurement device. The last experiment explores the concept of ‘accuracy’ in tooth colour measurement.

1.2 Tooth Colour

Dentition are polychromatic with a gradation of colour and shades (Vallittu *et al.*, 1996; Hammad, 2003). This colouring is the combined effect of the tooth’s structure and staining. Structurally, the semi-translucent enamel material does not fully obscure underlying dentine, making dentine the dominant contributor to the overall colour of the tooth (Joiner, 2004; Oguro *et al.*, 2016; Joiner and Luo, 2017; Brook *et al.*, 2007). Tooth colour changes over time due to either staining and/or aging. As individuals age, the enamel layer thins, increasing transparency and allowing for the underlying dentine to be more dominant as it continues to be formed, which significantly darkens or yellows the appearance of the tooth (Vallittu *et al.*, 1996; Schmidseder, 2000; Brook *et al.*, 2007; Joiner, 2004; Alkhatib *et al.,* 2005). Colour differs between tooth structures (i.e. canines are generally darker than incisors) and location on the tooth due to the fact that there is larger amount of dentine in the gingival area of the tooth than the incisal edge as well as the degree of translucency of the enamel varies widely between human teeth (i.e. gingival area is lightly darker than the middle and incisal edge) (Johnston, 2009; Brook *et al.*, 2007; Joiner, 2004; ASTM E2466-13, 2013).

Staining contributes to the appearance of teeth. Staining occurs when chromogens accumulate inside the tooth structure (intrinsic) or are adhered to the outside of the tooth structure (extrinsic) (Carey, 2014; Joiner, 2010). Intrinsic staining occurs either in the structure of the enamel

or dentine during tooth development or post tooth eruption and can be present on one tooth or generalised throughout the entire dentine (Joiner, 2010; Carey, 2014; Brook *et al.*, 2007). Different examples of discoloration are ageing (yellowing), excessive fluoride ingestion (white opaque spots or streaks to brown pitted patches), dental caries, severe jaundice in infancy, restorations, antibiotics such as tetracycline staining (blue-grey), white spot lesions, enamel microcracks, as well as genetic disorders such as amelogenesis imperfecta (disordered enamel formation causing cream, yellow to brown/black discoloration), dentinogensis imperfecta (abnormal dentine formation causing blue-grey to yellow brown discoloration), alkaptonuria (incomplete oxidation of phenylalaline and tyrosine causing accumulation of homogentisic acid which can cause brown discoloration), and enamel hypoplasia (Brook *et al.*, 2007; Alqahtani, 2014; Carey, 2014). This type of staining cannot be removed by tooth brushing or professional cleaning (Brook *et al.*, 2007).

Extrinsic stains, sometimes called external staining, occur due to the accumulation of pigmented materials into the plaque adhered to the outside of the tooth (Alqahtani, 2014; Joiner, 2010; Carey, 2014, Joiner *et al.*, 2008a). These types of stains are caused by poor oral hygiene/tooth brushing, smoking, or by dietary intake of certain foods such as red wine, coffee, and tea (Joiner *et al.*, 2008a). Sugars consumed in an individual's diet will interact with the bacteria within the plaque film on the outside of enamel to produce an acid which, left unmanaged, will slowly dissolve the enamel layer creating holes and dental caries (i.e. cavities which form yellow-brown staining) (Brook *et al.*, 2007; Mackie and Blinkhorn, 1995). Extrinsic stains can be prevented with a toothpaste with fluoride in it, can be removed or controlled by tooth brushing with abrasive toothpaste, removed by cleaning/removal of tartar by a dentist, or use of oral rinses (Pan and Westland, 2018; Joiner *et al.*, 2008a; Joiner, 2010; Joiner and Luo, 2017).

Tooth ‘whitening’ can be accomplished by physical removal of extrinsic stains, a chemical reaction to lighten the tooth colour, or a substance adhered to the tooth surface which changes its reflectance properties (Carey, 2014; Mohan *et al.*, 2008). Both extrinsic and intrinsic stains can be reduced or removed by bleaching (Joiner, 2010). Bleaching is defined as the chemical degradation of the chromogen (Carey, 2014). It is a chemical process in which hydrogen peroxide (or carbamide peroxide) diffuses into the enamel and dentine to decolourise or oxidise the chromogen (Joiner and Luo, 2017). The efficacy of the different bleaching methods is dependent on the type of tooth colour being treated, bleaching agent (hydrogen peroxide, carbamide peroxide, etc.), concentration frequency, duration of each application, and treatment period (Carey, 2014; Mohan *et al.*, 2008). The dental profession refers to two different types of bleaching: non-vital (intrinsic) and vital (extrinsic). Non-vital bleaching is for intrinsic stains in which the chemical is placed inside the pulp chamber. However, this method is extremely dangerous as it can cause dentine decay, also known as internal root resorption (Schmidseder, 2000). More common is vital tooth bleaching which is for extrinsic stains in which the chemicals are deposited onto the plaque and enamel to alter any discoloration (Schmidseder, 2000; Carey, 2014). Over-the-counter bleaching agents were first launched in the US in the late 1980’s (Alqahtani, 2014). Today, there are numerous different bleaching delivery formats available for home application such as toothpaste, strips, shields, trays, rinses or applicator brushes/paint on (Mohan *et al.*, 2008; Joiner, 2010, Carey, 2014). The most accessible product for consumers is whitening toothpastes (Joiner *et al.*, 2008a). With ‘whitening’ products, tooth colour can become 1-2 shades ‘whiter’ with the use of these products after 2 weeks of use; note that these at-home products typically have a lower concentration (3-6%) of peroxide (Carey, 2014; Joiner *et al.*, 2008a; Alqahtani, 2014). Even though it takes longer, home bleaching has become more common as it is more affordable, flexible when you want to bleach, and is safer due to the lower peroxide

concentration than going to the dentist for a bleaching treatment (Schmidseder, 2000). Dental Office bleaching treatments (i.e. trays, Therasmile) use higher concentrations of peroxide (15-38%) producing quicker colour changes (Joiner *et al.*, 2008a; He *et al.*, 2012; Schmidseder, 2000; Mohan *et al.*, 2008; Flucke, 2021). Bleaching can cause instability in the sealants, ceramic crowns, and composite restorations (Carey, 2014). An alternative method to make the teeth 'whiter' is the use of blue optical technologies (Joiner, 2010). Blue covarine in toothpaste is deposited onto the tooth surface to give a yellow to blue colour shift making the teeth appear perceptually less yellow (Westland *et al.*, 2017; Joiner and Luo, 2017). This is a temporary approach which can provide an instant perceivable change but only for a limited number of hours (Joiner *et al.*, 2008b).

1.3 Physics of Tooth Colour

Colour is the phenomenon in which visual perception responds to light when it is reflected or transmitted from an object (Kim-Pusateri *et al.*, 2009). Object colour has three components: light, the object and the observer. Light will interact with the object and some of that light is reflected, transmitted, or emitted to an observer, human or photodetector, which processes the light into a colour (Johnston, 2009; Kim-Pusateri *et al.*, 2009). Light interacts with the tooth in many ways such as specular and diffusion reflection at the surface and body of the tooth, absorption and scattering of light within the enamel and dentine, and transmission of light through the tooth exiting the back or edges of the material (Oguro *et al.*, 2016; Joiner and Luo, 2017; Burkinshaw, 2004) (see Figure 2).

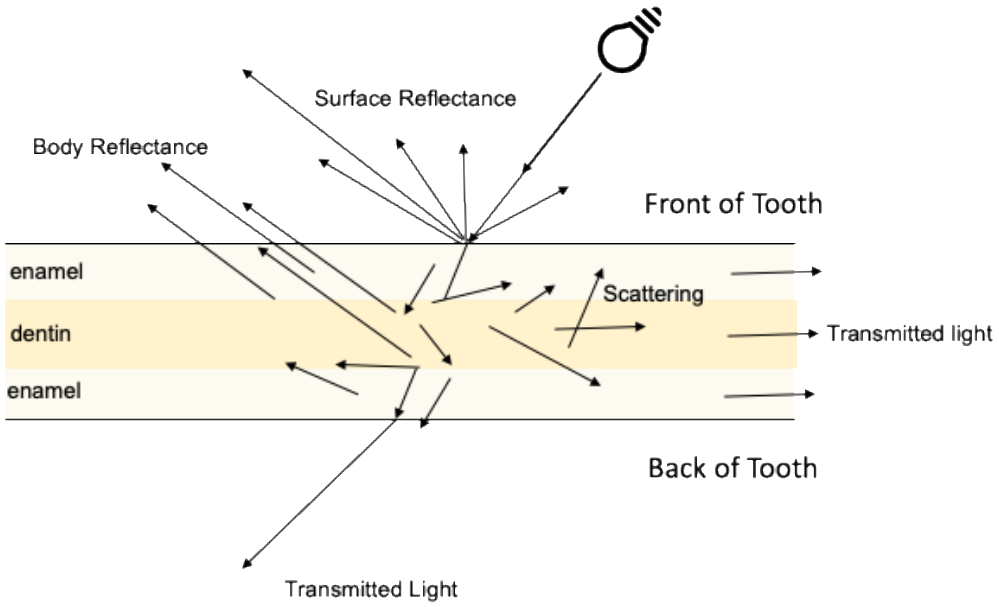

Figure 2: Schematic diagram of the different light interactions in layers of tooth

Absorbed light may excite electrons in the material producing fluorescence and emitted light (Nassau, 1998; Burkinshaw, 2004). Scattered light refers to the light waves changing directions within the subsurface of the material. Light that enters one boundary passing through a medium and exiting a different boundary is called transmitted light. Transmitted light occurs in transparent and translucent materials (i.e. glass, skin, teeth, soap) (Knight, 2008; Gevaux *et al.,* 2020). Reflected light refers to light which enters the subsurface of a material and quickly backscatters out the same side the light entered (Gevaux *et al.*, 2020). On the surface, both specular reflection and diffuse reflection can occur. Specular reflection is defined as the light that reflects at an equal but opposite angle, while diffuse reflection refers to light scattering in various directions (Figure 3) (Konica Minolta, 2019; Luo, 2011). Both diffuse and specular reflection can happen on the surface and body of the material.

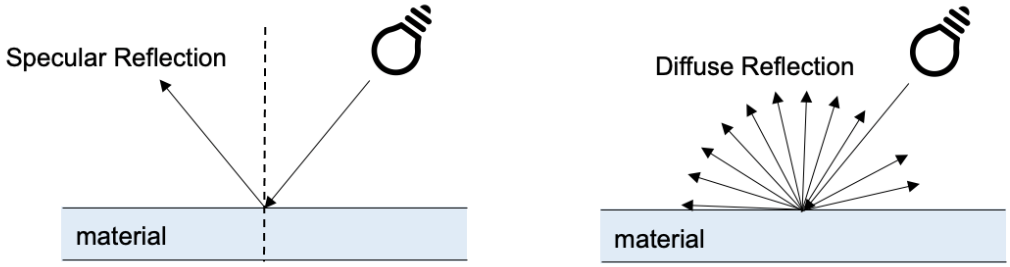

Figure 3: Schematic diagram of specular and diffuse reflection

Physically, tooth colour is a combination of light reflected from the enamel and underlying dentine that is observed and perceived (Brook *et al,* 2007). In addition, there is an amount of light that is scattered, absorbed, and transmitted light which is not seen. Light emerging from the tooth and reaching the eye is processed as a colour. Light that emerges from the tooth but which does not reach the eye (or which is unmeasured by the detector) is termed edge-loss (Joiner and Luo, 2017; Bolt *et al.,* 1994; Gevaux *et al.,* 2020). When more light is absorbed or transmitted, the tooth appears darker. When more light is reflected (surface and body) from the tooth, it will appear lighter (Brook *et al.*, 2007). Plaque, calculus, and saliva influence surface texture which changes the directionality of reflected light influencing overall tooth colour. A highly finished or smooth object will reflect more light specularly (Al-Azzawi, 2007), while a rough surface (i.e. plaque) scatters light more diffusely (Billmeyer and Saltzman, 1981; McDonald, 1997).

1.4 Colour Measurement

Colour measurement is the quantification of colour by measuring visible light that is reflected from, or transmitted through, an object visually or with an instrument. Clinically, visual assessment of tooth colour is still common practice despite being encumbered by numerous sources of errors (Bona *et al.*, 2009; Brook *et al.*, 2007). Objective colour measurement reduces and overcomes inconsistencies of traditional shade matching done by visual assessment such as metamerism (Lehmann *et al.*, 2010). Numerical colour data increases communication between industry

experts, accuracy of satisfactory colour matching, and the verification of restoration colour and efficacy of tooth whitening procedures (Lee *et al.*, 2011; Ahn and Lee, 2008; Chu *et al.*, 2010; Paravina, 2008; Beltrami *et al.*, 2014). In the last decades continual development of colour measurement devices have been used to evaluate tooth colour (Klotz *et al.*, 2018; Luo *et al.*, 2017; Lehmann *et al.*, 2010; Mohan *et al.*, 2008; Brandt *et al.*, 2017). Instruments such as colorimeters, reflectance spectrophotometers (SP), tele-spectroradiometers (SR), imaging systems (i.e. digital cameras), and combination instruments such as digital spectrophotometers have been used to measure tooth colour (Pan and Westland, 2018). Some instruments have been specifically designed for tooth colour measurement such as 3Shape TRIOS (digital camera and SP), CrystalEye (SP), Vita EasyShade (SP), ShadeVision (colorimeter), Degudent Shadepilot (colorimeter), SpectroShade Micro (digital camera and SP) (Da Silva *et al.*, 2008; Lehmann *et al.*, 2010). Some of these devices are no longer produced, but old instruments might still be used in practice. The development of instrumentation for tooth colour measurement aims to provide numerical colour data to increase communication of visual expectations and needs of dental restorations or whitening treatments quantitatively (de Bragança *et al.*, 2021; Da Silva *et al.*, 2008; Ahn and Lee, 2008: Tao *et al.*, 2017a; Collins *et al.*, 2008). These instruments provide numerical colour data with varying difficulty and accuracy (Pan and Westland, 2018). Teeth are polychromatic, small, non-planar, translucent and difficult to access in vivo. All these qualities make quantification of tooth colour difficult to analyse with instruments (Bona *et al.*, 2009; Lasseree *et al.*, 2011). It is unclear what the most accurate measurement method is for quantifying tooth colour, but there are advantages and disadvantages to each method (Pan and Westland, 2018). Currently, it is suggested that using both visual and instrument methods together for measuring tooth colour and colour matching may provide more predictable aesthetic outcomes as there is no ‘gold standard’

instrument or methodology for tooth colour measurement (Chu *et al.*, 2010).

1.4.1 Visual Assessment

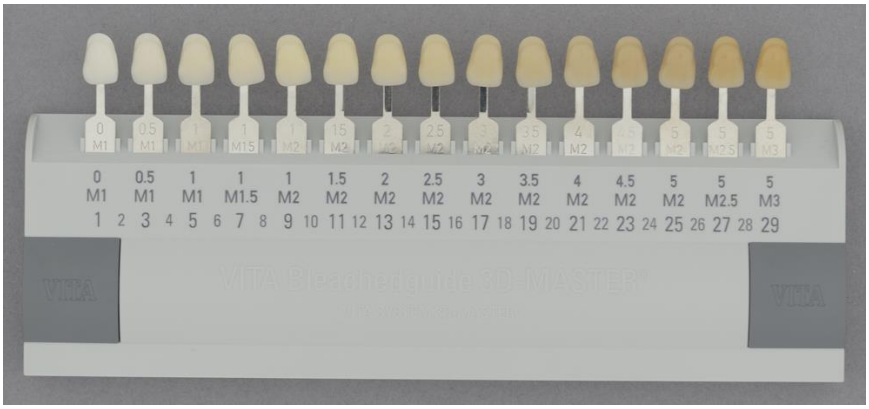

Figure 4: Vita Bleachedguide 3D-Master shade guide

Figure 4 displays an example of a typical shade guide on the market. Shade guides are sets of tooth shaped porcelain or ceramic tabs differing slightly in lightness, chroma and hue from one another that are similar to human teeth in size, shape and structure increasing chances for successful shade matching (ASTM E2466-13, 2013; Kim-Pusateri *et al.*, 2009; Chu *et al.*, 2010). Visual evaluation is performed by a human comparing a porcelain tooth shade tab from a standardised shade guide with a human tooth and determining the closest match for the evaluation of a whitening treatment pre and post treatment or for the match to be communicated to another dental professional to create a restoration (Paravina, 2008; Hammad, 2003; Kovacevic Pavicic *et al.*, 2017). Shade guides available on the market include Vitapan Classical, Vitapan 3D Master, Vita Lumin Vacuum guide, Vitapan Tooth Guide, Vitapan Bleached guide, Chromascop and Vitapan Linearguide (Chu *et al.*, 2010; Ahn and Lee, 2008). They are intended to cover the entire tooth colour gamut (Brook *et al.*, 2007). However, the majority of shade guides do not cover the full range of tooth colour, lack logical distribution between tabs, and may differ between manufacturers although differences may even occur in guides made by the same manufacturer (Joiner and Luo, 2017;

Khurana *et al.*, 2007; Da Silva *et al.*, 2008; Lasseree *et al.*, 2011; Lehmann *et al.*, 2010; Tung *et al.*, 2002; Curd *et al.*, 2006; Ahn and Lee, 2008). While the dental industry acknowledges that shade guides lack standardisation and colour uniformity, visual assessment still remains the oldest and most frequently used method for assessing and communicating colour in clinical dentistry (Paravina, 2009; Mohan *et al.*, 2008; Khurana *et al.*, 2007). This might be due to the fact that it is a very quick, easily available, and a cost effective method for clinical use (Joiner, 2004; Kovacevic Pavicic *et al.*, 2017).

The human visual system is sensitive and can detect small colour differences (Kovacevic Pavicic *et al.*, 2017; Chu *et al.*, 2010). However, visual colour determination is not consistent between different clinicians; even highly trained observers inevitably make wrong decisions 20% of the time and increases when the tolerance tightens (McLaren, 1970; Khurana *et al.*, 2007). This is due to inconsistencies between human observers such as gender, fatigue of the eye and colour vision deficiencies caused by birth, age, diseases or drugs (Joiner *et al.*, 2008a; Kovacevic Pavicic *et al.*, 2017; Curd *et al.*, 2006; Gómez-Polo *et al.*, 2015; Chu *et al.*, 2010). For teeth especially, differences in understanding and perceiving colour, experience of using shade guides, lighting conditions, and the effects of surrounding gingival colour can influence shade selection (Joiner and Luo, 2017). The eye could also be distracted by the gloss, translucency, surface irregularities and shape of the tooth (Lasseree *et al.*, 2011). The use of a shade guide is an estimation of tooth colour and many argue that this method of tooth colour analysis is considered too subjective and prone to error (Brook *et al.*, 2007). Instrumental colour measurement is often preferred (Pop-Ciutrila *et al.*, 2016).

1.4.2 Instrumental Colour Measurement

Dental measurement devices can be categorised as contact or non-contact based devices. Contact devices, such as spectrophotometers and colorimeters, make direct contact with the surface of the object. Non-

contact based devices, such as spectroradiometers and digital cameras, do not physically make contact with the object. Some categorise devices based on either complete tooth measurement devices or spot measurement. Spot measurement devices do not take colour measurement of the entire tooth but rather a ‘spot’ or small portion of the tooth. Complete tooth measurement devices measure the entire area of the face of the tooth surface (Khurana *et al.*, 2007).

Spectroradiometers, colorimeters and spectrophotometers are spot or limited-area measurement devices meaning the sensor only captures about 1-5mm diameter areas of the tooth (Chu *et al.*, 2010). The results from spot measurement may be unrepresentative of the entire tooth colour as they are limited to the aperture size (Brook *et al.*, 2007; Pan and Westland, 2018). Teeth are non-homogeneous in shade and even minor surface variations occur due to abrasions or impurities may influence colorimetric measurements (Lasseree *et al.*, 2011; Seghi, 1990). Some teeth have ‘darker’ incisal edges, which might not be picked up by a spot measurement devices leading to an unrepresentative measurement of the whole tooth surface (Guan *et al.*, 2005; Khurana *et al.*, 2007). However, three areas of the tooth are usually measured in order to obtain a ‘representative’ evaluation of tooth shade: gingival, central and incisal (Chu *et al.*, 2010). Some suggest that the free hand positioning of these contact based/spot measurement devices causes suboptimal repeatability and is more prone to measurement errors due to the tip’s positioning, angle, contact on the tooth and movement from patient or operator (Lasseree *et al.*, 2011; Chu *et al.*, 2010; Bona *et al.,* 2009). Digital cameras or hybrid colour measurement devices that include a digital camera such as the 3Shape TRIOS (digital camera and SP) and the SpectroShade Micro (digital camera and SP) can be complete tooth colour measurement devices (Da Silva *et al.*, 2008; Lehmann *et al.*, 2010). They present an image of the whole tooth in one image and the colour data from all the pixels is amalgamated (Khurana *et al.,* 2007; Chu *et al.,* 2010).

Colorimeters are contact-based and spot measurement devices which measure tristimulus values (see section 1.5.1) by filtering reflected light from an object into red, green, and blue areas of the visible spectrum (Joiner and Luo, 2017; Brook *et al.*, 2007; Kim-Pusateri *et al.*, 2009; Tung *et al.*, 2002; Chu *et al.*, 2010; Johnston, 2009). The advantages of this device are that it provides results in terms of colour space, is easy to use, is sensitive in detecting and measuring small colour differences, and less expensive than other devices such as spectrophotometers (Brook *et al.*, 2007; Joiner and Luo, 2017; Kim-Pusateri *et al.*, 2009). However, over time the filters will age and this increases measurement errors (Kim-Pusateri *et al.*, 2009).

Spectrophotometers are contact-based devices that measure the light reflected (surface and body) from a stimulus compared to a standardised stimulus at each wavelength along the visible spectrum (often from 360nm to 780nm and at 10nm intervals) providing a ratio of the two amounts called the reflectance factor (Brook *et al.*, 2007; Joiner and Luo, 2017; Kim-Pusateri *et al.*, 2009; Chu *et al.*, 2010; Tung *et al.,* 2002). The reflectance factor can be converted into useful colorimetric values for any CIE standard illuminant (VITA Zahnfabrik, 2013; Palumbo and Weber, 1999). Due to the device's ability to measure the amount of light reflected from an object over the full spectral wavelength, it is considered better than colorimeters as it provides more systematic and precise measurements (Da Silva *et al.*, 2008; Johnston, 2009). Widespread use of these devices is hindered due to their complexity (require training to use) and cost (Khurana *et al.,* 2007; Joiner, 2004; Kovacevic Pavicic *et al.*, 2017; Curd *et al.*, 2006).

Disadvantages of contact-based devices include the difficulty of *in vivo* measurements, potential cross-infection, fogging of optical lens, tip making full contact with a curved surface, as well as gingival discomfort

caused by the instrument tip (Khurana *et al.*, 2007; Pop-Ciutrila *et al.,* 2016; Da Silva *et al.,* 2008; Tung *et al.,* 2002; Brook *et al.*, 2007; Joiner and Luo, 2017). Dental specific contact-based colour measurement devices eliminate these problems (Tung *et al.,* 2002). However, it has been suggested that contact-based devices are prone to significant edge-loss due to the small window size of the aperture (Bolt *et al.,* 1994). Non-contact based devices (i.e. spectroradiometer and digital cameras) also eliminate in vivo problems. Some researchers suggest edge-loss is reduced with non-contact based devices as the aperture does not restrict external light sources or ambient lighting (Lee *et al.,* 2011; Pop-Ciutrila *et al.,* 2016; Guan *et al.,* 2005; Joiner and Luo, 2017; Bolt *et al.,* 1994). Edge-loss is inevitable with translucent materials (i.e. teeth); the amount of edge-loss is device dependent.

Spectroradiometers are non-contact based spot measurement devices which measure radiometric quantities (i.e. irradiance and radiance) that is emitted or reflected from an object (Joiner and Luo, 2017; Pan and Westland, 2018). The main advantage to these instruments is that issues regarding curved surfaces and edge-loss are minimized. However, they are expensive, require an external light source, and need training to ensure correct use (Guan *et al.,* 2005; Pan and Westland, 2018; Lee *et al.,* 2011).

Digital cameras are non-contact based systems that record reflected light onto a light sensing material and output an image represented by red, green and blue values (RGB) at each pixel (Joiner and Luo, 2017; Chu *et al.,* 2010). RGB values are device dependent meaning that the RGB values for the same object could change for different cameras (Wee *et al.*, 2006). Regardless of the device, RGB values are arranged to produce an image that looks like a topographical colour map of the whole tooth (Chu *et al.*, 2010; Khurana *et al.*, 2007; Wee *et al.,* 2006). Mathematical transformations can be used to convert RGB values of each pixel into CIE colour coordinates (Joiner and Luo, 2017; Guan *et al.*, 2005). A digital

image thus provides colour information at each spatial position on a tooth (spot measurement) or provide a complete tooth measurement unlike spectrophotometers, colorimeters, or spectroradiometers (Pan and Westland, 2018; Brook *et al.,* 2007; Mohan *et al.,* 2008). Digital cameras are sometimes preferred in clinical situations for this reason (Guan *et al.*, 2005). The advantage of digital images is that they can provide visual evidence of efficacy of aesthetic treatments (i.e. before and after images), produces documentation of injuries for insurance companies, as well as produce a database of images archiving health records to be analysed or reanalysed at a later date (Schmidseder, 2000; Brook et al., 2007; Pan and Westland, 2018; Mohan *et al.,* 2008). Digital cameras are readily available and relatively cheap (Wee *et al.,* 2006). However, image quality is a critical factor which requires expertise for proper calibration which includes adjusting focal length, F-stop, and shutter speed appropriately as well as adjusting the white balance. (Wee *et al.,* 2006; Joiner and Luo, 2017; Pan and Westland, 2018; Chu *et al.,* 2008, Da Silva *et al.,* 2008). Some digital cameras also require additional equipment such as external standardised light source which are expensive and some suggest the addition of a polarising filter in order to exclude specular reflection (Joiner and Luo, 2018; Guan *et al.,* 2005).

No matter what device is used, colour measurement of an object is influenced by the angle in which the object in illuminated and angle in which light is collected (measurement geometry), illumination used, contact or non-contact based devices, and distance which object is observed (standard observer) (Joiner *et al.,* 2008a).

1.4.2.1 Measurement Geometries

Measurement geometry refers to the angle at which a light source illuminates a sample in relation to the angle at which reflected or transmitted light is detected/viewed in relation to where the sample is (Hunt and Pointer, 2011). It is important to define the measurement

geometry when taking colour measurement as the illumination angle directly affects how light is reflected from an object. A sample may appear brighter, duller, lighter or darker depending on the angle at which it is viewed (Konica Minolta Sensing Americas Inc., 2020). Reflectance spectrophotometers, gloss meters, and colorimeters have an internal optical geometry system, which means that a source of illumination and detector is built into the device. Other systems such as tele-spectroradiometers measuring non-emitting light samples or the use of a digital camera for colour measurement require additional equipment to create measurement geometries such as a copy-stand, controlled viewing cabinet, or controlled light box. CIE standardised measurement geometries for colour measurement. While most devices follow a standardised geometry, not all devices do such as Vita EasyShade.

The Commission de L'Éclairage is an organization devoted to the standardisation of colour measurement by defining standard light sources, observers, instrumentation, measurement geometries, and provides methods for converting measured light into useful colorimetric data (Joiner, 2004). There are different CIE standardised measurement geometries: bidirectional, multi-directional and diffuse (ASTM E1349-06, 2018). Choice of measurement geometry is dependent on material being measured as well as desired information. A multi-angle spectrophotometers or gloss-meters are designed to provide information about colour of the sample at different angles as if it was moved back and forth. It is used for the measurement of gonio-apparent materials such as metallics, pearlescence, mica, plastics, effect paints, etc (ASTM E179-17, 2017; Konica Minolta, 2016b). The illumination is set at a 45° angle from the sample with 6 different angle observations: -15°, 15°, 25°, 45°, 75°, and 110° (Figure 5) (Konica Minolta, 2016b).

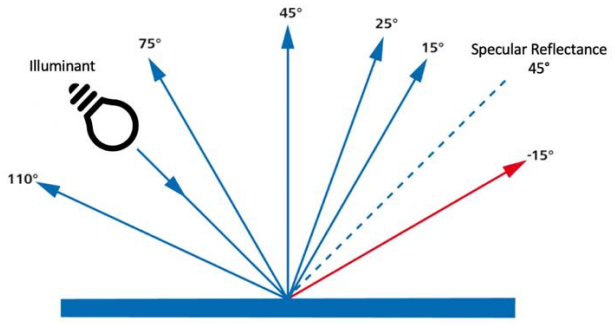

Figure 5: Diagrammatic representation of multi-angle measurement geometry

A bi-directional geometry (45°/0° or 0°/45°) is recommended for measuring retroreflective materials, paints, papers, porcelain enamels, textiles, non-metallic surfaces, fluorescent, and objects with intermediate gloss (ASTM E1349-06, 2018; ASTM E179-17, 2017). With a 0°/45° system, the illumination occurs at a 0° angle from the sample and the detector records at a 45° angle (Figure 6). This geometry can be achieved using a lighting cabinet and positioning a device (i.e. tele-spectroradiometer, digital camera) at a 45° angle.

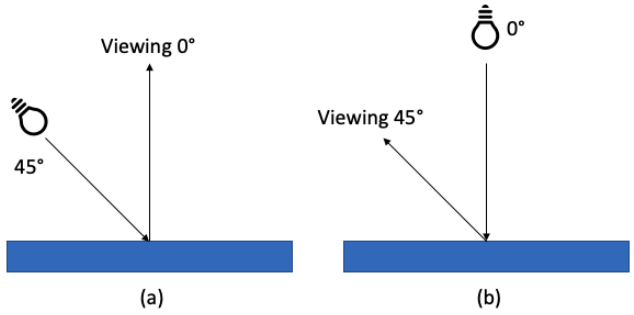

Figure 6: Schematic diagram of measurement geometries (a) 45°/0° and (b) 0°/45°

For 45°/0°, one or more illumination beams are positioned at an 45° angle in relation to the sample with the detector positioned at a 0° angle above specimen (Figure 6). Commercial reflectance spectrophotometers are available with this geometry, with a light source built into the device. According to the American Society for Testing and Materials (ASTM)

E179-17 (2017) a 45/0° geometry should be used for results that relate best to visual assessment and assess appearance attributes of an object. While this geometry can be achieved for the digital camera and tele-spectroradiometer combined with a lighting cabinet or copy-stand, there are problems with 45°/0° and 0°/45° reflectance spectrophotometers. Teeth can have a diameter anywhere from 4.5mm to 10mm (Chu, 2007; Chu and Okubo, 2008; German *et al.*, 2016). The majority of commercial bi-directional spectrophotometers' apertures are too large for accurate colour measurement of teeth.

Diffuse/0° or 0°/diffuse geometry is based on viewing in a completely diffuse illumination. Diffuse illumination is when light is distributed in all directions (ASTM E179-17, 2017). Lighting boxes and lighting cabinets can provide diffuse illumination for devices without internal light source such as digital cameras and tele-spectroradiometers. Devices with internal light sources such as spectrophotometers have a hemispherical design to produce diffuse illumination.

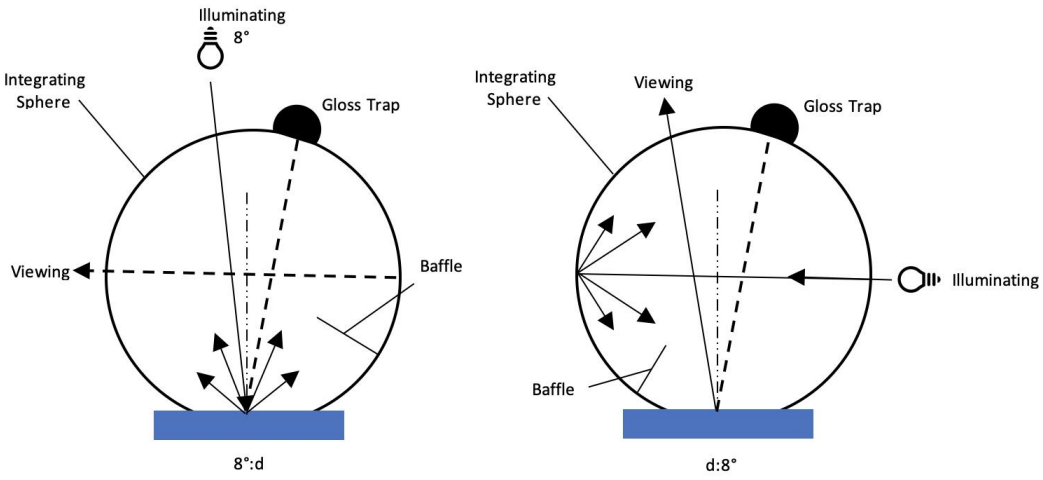

Figure 7: Diagrammatic representation of CIE hemispherical geometries

However, the collection sensor occurs at an 8° $\pm$ 2° angle instead of at 0° (ASTM E2466-13, 2013) (Figure 7). Receiver or illumination beams may have an angle up to 10° to be compliant with existing CIE standards

(Palumbo and Weber, 1999; ASTM E179-17, 2017). This hemispherical design allows for spectrophotometer devices to measure a sample with or without including the spectral reflection.

Specular component included (SCI) means device measures both specular reflected and diffuse reflected light. Specular component excluded means the measurement excludes any specular reflected light (i.e. specular surface reflectance, specular body reflectance). By definition, bidirectional geometry devices (45°/0° or 0°/45°) as well as perfect diffuse/0° geometry collect specular component excluded measurement. The allowance of a measurement angle up to 10° allows for the specular component to be included in measurements (Palumbo and Weber, 1999; ASTM E179-17, 2017). Devices with external light sources can be set up with an angled geometry such as diffuse/angled (i.e. 5°,10°, 15°, 45°) in order to produce specular included measurements. It is unclear why the CIE has determined that 10° is the max angle for it to be compliant with their standards.

1.5 Commission de L'Éclairage (CIE)

The CIE have developed the most widely used systems for describing colour (Tung *et al.,* 2002). Data collected from various devices are mathematically translated into useful colorimetric data, which can be presented in different colour spaces: most notably CIEXYZ (also known as tristimulus values), chromaticity coordinates (xyz), or CIELAB (Chu *et al.,* 2010).

1.5.1 Tristimulus Values (XYZ)

In 1931, the CIE developed a method for calculating tristimulus values, which are three numbers that represent how the visual system responds to a spectral stimulus (Pan and Westland, 2018). CIE used a system of imaginary primaries, called CIE XYZ, so that they would always have positive values (Gupte, 2010). Spectral data can be converted to CIEXYZ

values using the following Equations 1-1 to 1-3 (Wyszecki and Stiles,1967; CIE 15, 2004).

$$\boldsymbol{X = k\sum_{\lambda=360}^{\lambda=740} E(\lambda)\bar{x}(\lambda)P(\lambda)}$$

$$\boldsymbol{Y = k\sum_{\lambda=360}^{\lambda=740} E(\lambda)\bar{y}(\lambda)P(\lambda)} \qquad \textbf{Equation 1-1}$$

$$\boldsymbol{Z = k\sum_{\lambda=360}^{\lambda=740} E(\lambda)\bar{z}(\lambda)P(\lambda)}$$

where

$$\boldsymbol{k = 100/\int_{\lambda}^{\lambda} E(\lambda)\bar{y}(\lambda)d\lambda.}$$

E(λ) is the relative power distribution of the light source, P(λ) is the spectral reflectance of the object, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are the colour matching functions for the CIE Standard Observer, and k is a constant in which normalises the tristimulus values based on the perfect reflecting diffuser (Y=100). Weighing tables are where the terms E(λ)$\bar{x}(\lambda)$, E(λ) $\bar{y}(\lambda)$, and E(λ) $\bar{z}(\lambda)$ are pre-calculated at each wavelength with the normal constant k factored into it. The weighing factors are calculated from colour matching functions based on a standard observer and standard illuminant used.

1.5.1.1 Standard Observer

The Standard Observer is the average response of a group of individuals and is defined for two so-called visual angles, which relates to the size of an object and the distance at which it is viewed (Hunt and Pointer, 2011). The CIE has defined two standard observers, 2° and 10°; each define the field of view for which the colour matching values of an ideal observer correspond to CIE colour-matching functions for that field size (Gupte, 2010). In 1931, the CIE determined that a 2° viewing angle of an object would result in most of the colours of an object hitting the fovea of the eye, an area of the eye which holds the most colour-detecting cones (Konica Minolta, 2015; Hunt and Pointer 2011; Nassau, 1998).

In 1964, the CIE conducted another set of experiments with a 10° field of view known as the 10° observer. This is more commonly used today in

fields such as textiles as it closely approximates industrial colour matching and quality control viewing conditions (Nassau, 1998). With a viewing distance of 50cm, a 2° field of view would observe approximately a 1.7 cm diameter circle of an object or a dime at arm's length. While a 10° field of view would view approximately a 8.8 cm diameter circle of an object at a 50cm distance, Figure 8 (Konica Minolta, 2015; Nassau, 1998; CIE 15, 2004).

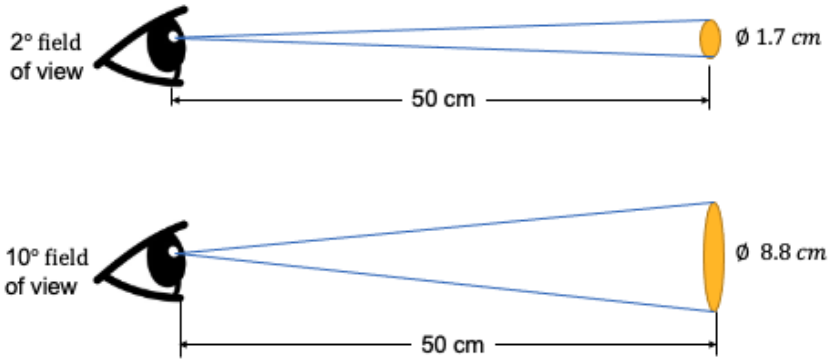

Figure 8: Comparison of CIE Standard Observers 2° and 10°

The largest samples used in this thesis have a 2 cm diameter and the farthest distance in which an instrument's sensor is from any sample is 100 cm. This would create a field of view between 1° and 4°, therefore a 2° standard observer is used for data conversion in this thesis (Konica Minolta, 2015; Nassau, 1998). If the viewing angle was above 4°, a 10° standard observer would be used for calculations (Konica Minolta, 2015).

1.5.2 Chromaticity Coordinates (xyz)

Chromaticity coordinates are the ratio of the individual tristimulus values and the sum of all three tristimulus values. The chromaticity coordinates are denoted by x, y, z. The chromaticity coordinates can be calculated from XYZ tristimulus values using the following equations (CIE 15, 2004; Hunt and Pointer, 2011):

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

$$x = \frac{X}{X+Y+Z} \qquad \textit{Equation 1-2}$$

$$z = \frac{Z}{X+Y+Z}$$

Only two coordinates are needed to specify chromaticity as the sum of x, y, and z, will always be equal to one.

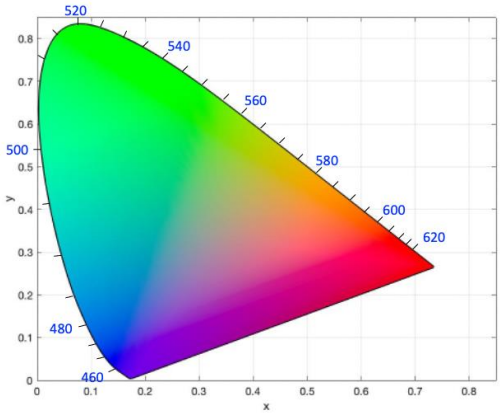

Figure 9: The CIE chromaticity diagram with spectrum locus energy values (nm)

A chromaticity diagram can be plotted using x and y values (Figure 9). The curved line of the diagram is called the spectrum locus which represents monochromatic stimuli throughout the spectrum indicated by nanometres or spectrally pure colours (Hunt and Pointer, 2011; Malacara, 2002). The straight line is known as the purple boundary which is a presentation of the combination of red and blue in different proportions. The limitation of this colour space is that only saturation and hue are represented in the

chromaticity diagram. When x=0.33 and y=0.33 that is called the equi-energy stimulus, which represents white, black and all grey values. Increasing saturation of a sample is represented by the values moving towards the spectrum locus (Hunt and Pointer, 2011; Malacara, 2002). For full classification of a colour, luminosity must be included which is why Y from CIEXYZ is normally cited with chromaticity coordinates x and y.

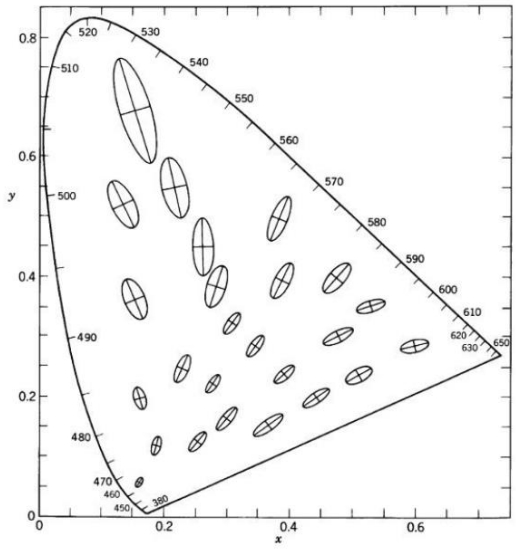

Figure 10: MacAdam Ellipses showing non-uniformity of the chromaticity diagram **(Kuehi, 2003)**

Distances between two colours on a chromaticity diagram do not correspond to the degree of colour difference. It is expected that the ellipses in Figure 10 would be circles and have the same size if it was a uniform colour space. The non-uniformity of the chromaticity diagram is confirmed by MacAdam ellipses, in which the perceived colour differences are less towards upper part of the diagram (i.e. green area) versus the lower part of the diagram (MacAdam, 1942).

1.5.3 CIELAB (LAB)

In 1976, the CIE defined uniform colour space called CIE (1976) L*a*b*, also known as CIELAB (Gupte, 2010; Joiner, 2004). This colour space is the most frequently used in both in vivo and in vitro research for the colour characterisation of dental materials and teeth (Perez *et al.*, 2016). CIELAB

colour space was developed to correspond better to perceived colour differences (Ebner, 2007; Westland *et al.*, 2012). Three coordinates are used to define a colour: L*, a* and b* (Konica Minolta, 2015).

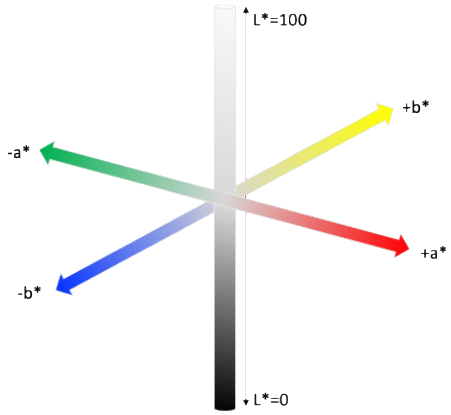

Figure 11: CIELAB colour space diagram

L* represents lightness (100) to darkness (0), a* represents the red-green axis, b* represents the blue-yellow axis (Figure 11). For a perfect white L*=100, and a perfect black L*=0, with a* and b*=0. A positive a* value corresponding to the colour being red or in red direction and negative a* value corresponds to the colour being green or in the green direction. Blue or a colour being in the blue direction corresponding to negative b* value, while yellow colour or in the yellow direction corresponding to positive b* value (Oleari, 2016; Gupte, 2010; Ebner, 2007). CIELAB values can be calculated from XYZ values using equations:

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16$$

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

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

where

$$f(I) = I^{1/3} \quad \text{if} \quad I > \left(\frac{6}{29}\right)^3$$

and

$$f(I) = \left(\frac{841}{108}\right)(I) + \frac{4}{29} \quad \text{if} \quad I < \left(\frac{6}{29}\right)^3$$

Variables X_n, Y_n, and Z_n are the tristimulus values for the chosen reference white. These values are usually based on a perfectly reflecting diffuser to which the illuminant has Y_n =100 (Gupte, 2010; Westland *et al.*, 2012).

1.6 Tooth Colour Gamut

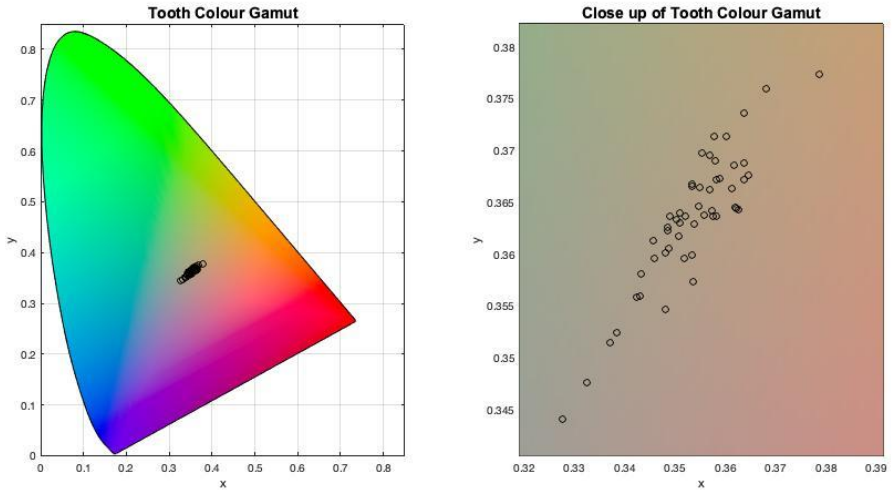

Figure 12 Tooth colour gamut placed on a chromaticity diagram on the left and a close up of the range of the tooth colour gamut on the right.

Tooth colour gamut is small in comparison the entire colour gamut.

There is no range of chromaticity values noted in studies for tooth colour. In this thesis the range for chromaticity values are x=0.3249-0.3939 and y=0.3397–0.3871 as shown in Figure 12.

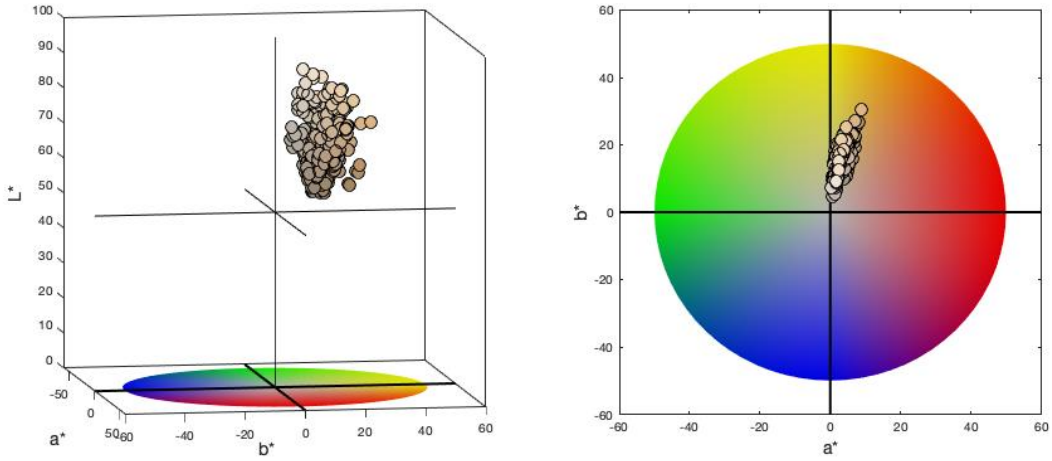

Figure 13: Tooth colour gamut on CIELAB plot

CIELAB is more commonly used in dentistry. Figure 13 displays the limited tooth colour gamut compared to the entire colour gamut in CIELAB space. Tooth colour gamut in CIELAB colour space has L* values of 55 to 95, a* values between 3 and 12 and b* from 8 to 25 (ASTM E2466-13, 2013). While this standardisation was withdrawn, their CIELAB range for tooth colour is within reason in comparison to other studies. Depending on the instrument and sample selection, the tooth colour gamut varies. Tooth colour gamut according to data collected by colorimeters produce L*= 30-85.6, a*= -6.1-9.8 and b*= -1-48.6 (Cho *et al.,* 2007; Yamanel *et al.*, 2010). Spectrophotometers have been reported to produce a range of L*= 48-89.6, a*=-4.5-7.3 and b*=-6-38.9 (Jarad *et al.,* 2005; Mahn *et al.,* 2021; O'Brien *et al.,* 1990; Yuan *et al.,* 2007; Paravina *et al.,* 2006). While, digital cameras have reported tooth colour to be from L*=52-92, a*= -4.5-2.5 and b*=2-25.5 (Yamanel *et al.,* 2010; Mahn *et al.*, 2021; Jarad *et al.,* 2005).

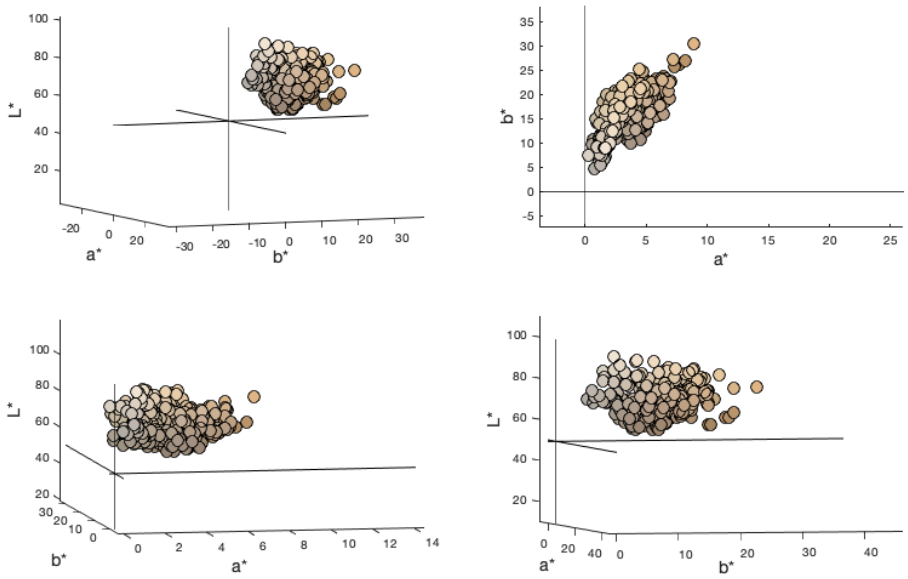

Figure 14: Tooth colour gamut according to measurements taken in this thesis

In this thesis, the range of tooth colour gamut was found to be closer to the range made by ASTM E2466-13 (2013), with L* values between 56 and 91, a* values ranging from 0.24 to 8.8 and b* values ranging from 4.8 to 36.6. Tooth colour range in CIELAB space is shown in Figure 14. It should be mentioned that it is a specific combination of CIELAB values that represents a tooth colour and not all combinations of values that fall within the tooth colour gamut are representative of teeth.

1.7 Whiteness and Yellowness Indices

Indices are equations that aid in computing the quantification of perceptual changes using colorimetric data. In dentistry, it is more common to use whiteness indices and yellowness indices to evaluate perceptual changes in tooth colour than CIE colour difference formulae (Pan *et al.*, 2018, Westland *et al.*, 2017).

A whiteness index computes the degree of departure of an object from that of a preferred white. In 1981, the CIE developed a whiteness index, WIC (Equation 1-4), to address the needs of paints, ceramics, textiles and

plastics based on CIE 1931 Yxy colour space (Joiner *et al.*, 2008a; Pan *et al.,* 2018; ASTM E313-05, 2005).

$$WIC = Y + 800(x_n - x) + 1700(y_n - y) \qquad \textit{Equation 1-4}$$

This index has been used in dental research for evaluating whiteness; however, it may not be useful in dentistry as it was developed for use on different materials (Perez *et al.,* 2016; Joiner and Luo, 2017). Two equations have been optimised to assess tooth whiteness: WIO and WID. WIO has the same format as WIC. Luo *et al.* (2009) developed WIO (Equation 1-5) by optimising the coefficients to best fit experimental data on the perception of tooth whiteness.

$$WIO = Y + 1075.012(x_n - x) + 145.516(y_n - y) \qquad \textit{Equation 1-5}$$

More recently Perez *et al.* (2016) developed a WI_D (Equation 1-6), which is an index based on the CIELAB colour space.

$$WI_D = 0.51L^* - 2.24a^* - 1.100b^* \qquad \textit{Equation 1-6}$$

In industry, a white colour departs from a perfect white in two directions either towards yellow or towards grey (Joiner and Luo, 2017). Whiteness has sometimes been expressed by the quantification of yellowness (Luo *et al.,* 2009). A yellowness index calculates the degree of departure of an object from a preferred white towards the yellow hue (ASTM E313-05, 2005; Hunter, 1981). There are at least 25 different yellowness indices but only two have been used in dental research: YIE313 and YID1925 (Equations 1-7 an 1-8) (ASTM E313-05, 2005, Hunter, 1981). These were designed for use in textiles, paints, oils and plastics meaning they might not be suitable to tooth colour analysis (ASTM E313-05, 2005).

$$YIE313 = 100\left(1 - \frac{0.847Z}{Y}\right) \qquad \textit{Equation 1-7}$$

$$YID\ 1925 = \frac{100(1.28X - 1.06Z)}{Y} \qquad \textit{Equation 1-8}$$

2 Methodology

2.1 Introduction

This chapter describes all the different colour measurement devices, psychophysical methods, data conversion equations, calculations, and data analysis techniques used in this thesis. Specifically, details are provided for the instruments used, their required warm up time, and the settings used. The different psychophysical methods are discussed, explaining the benefit of the ranking method that was used in this thesis. In addition, this chapter details how to convert that subjective data into useful objective values as well as the calculations used to convert the different physical measurement data into useful colorimetric values to be used for statistical analysis.

2.2 Instruments

It is standard practice for three colour measurements of each tooth to be collected: one measurement each at the gingival, central, and incisal locations of the tooth (Da Silva *et al.*, 2008). To replicate this procedure, in this thesis three measurements were taken of each sample by each device and averaged to produce one set of data for each sample from each device. The settings for each devices are provided below.

2.2.1 Konica Minolta CM-2600d Reflectance Spectrophotometer

The Konica Minolta CM-2600d is a sphere based diffuse/8° reflectance spectrophotometer collecting spectral data from 360nm to 780nm at 10nm intervals. This specific spectrophotometer is a contact-based device and provides its own illumination.

The first step is to turn on the reflectance spectrophotometer and allowing the device to reach temporal stability. The time in which a device reaches temporal stability will vary depending on the device. The temporal stability of this device was assessed in order to determine the time required for the

device to warm up for more precise measurements. After calibration, measurements of the white calibration plate (CM-A145) provided with the instrument were taken every 30 seconds for 4.5 hours (4.5 hours was chosen to provide ample warm-up time).

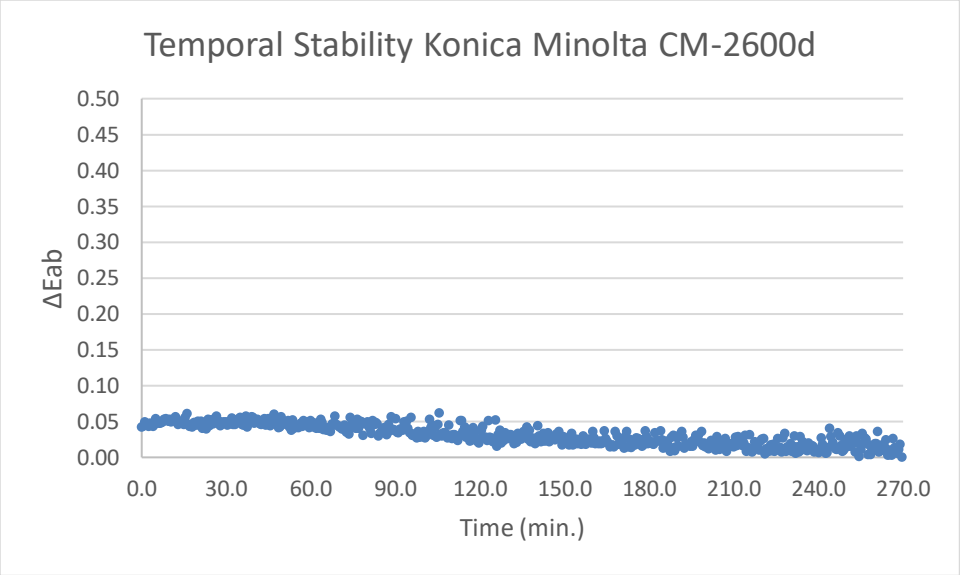

Figure 15: Temporal Stability of Konica Minolta CM-2600d over the course of 270 minutes (colour difference is computed by comparing each sample with the very last-measured sample)

The colour difference, ΔEab, was calculated by computing the difference between each measurement and the very last measurement (after 4.5 hours). Theoretically, the device reaches temporal stability when ΔEab reaches 0. However, this will only happen with the last measurement as the device inherently has measurement variability. According to Konica Minolta (2014), this device has inter-instrument variability of ΔEab of 0.2 after the device has been calibrated (Konica Minolta, 2014). As seen in Figure 15, the device is stable immediately after calibration with minor variation in ΔEab of less than 0.1. The largest ΔEab between measurements is 0.06 throughout the entire 4.5 hours indicating that the device will produce reliable colour measurements for the duration of experiments conducted in this thesis. Nevertheless, for all experiments where the Konica Minolta CM-2600d was used in this thesis, the device

was allotted 30 minutes to warm up where it produced a ΔEab value of 0.04.

In remote mode the device can be connected to a PC and controlled by SpectraMagic NX software. The remote setting was chosen for all experiments conducted in this thesis. Measurement conditions include choosing a measurement area (target mask attachment), specular component mode, standard illuminant, standard observer, UV, gloss, display mode, colour space, manual averaging, pass/fail, and delay time (Konica Minolta, 2014). Target mask attachment is used to change the specimen measuring port size which changes the illumination area and measurement area, which is chosen depending on the size of the sample. This setting needs to be changed via a switch on the side of the device as well as on the software. Due to the small size of the samples, SAV (small aperture size), which has an aperture of 6mm diameter (illumination area) and a 3 mm measurement area was used. The hemispherical design allows the device to measure specular included, specular excluded, or simultaneous measurements of specular included and specular excluded (Konica Minolta, 2014). Both components are measured in thesis. Standard illuminant, standard observer, and colour system measurement conditions provide the device and software information in order to immediately compute useful colorimetric data from spectral data collected (ASTM E1349-06, 2018; Konica Minolta, 2016a; Konica Minolta, 2014). While the software and device are capable of doing this conversion, in this thesis the colorimetric data was calculated from the spectral data without the software which is explained in section 2.6 so that the standard observer chosen and illuminant were irrelevant.

Once all device settings are chosen, there is a two-part calibration: a zero calibration and a white calibration. A zero calibration measures the amount of stray light in advance in order to compensate for the effects of stray light on each measurement. A white calibration, which occurs immediately after

the zero calibration, is the process in which reflectance of a white known calibration plate is measured to provide a reflectance scale. This white plate is usually provided with the device.

To measure a sample, the measurement port is placed in contact with the sample (ASTM E1349-06, 2018); a view finder is slid open to ensure positioning. The measurement button is then pressed allowing the 3-pulse xenon lamp to flash and measurement of the reflected light to be collected to provide spectral reflectance factors from 360 to 780nm at intervals of 10nm. For each measurement taken, two sets of data are produced: one specular included and one specular excluded.

2.2.2 X-rite 962 Spectrophotometer

The X-rite 962 is a reflectance spectrophotometer that measures reflectance factors from 360nm to 750nm at a 10nm intervals (X-rite, 2021). This instrument is a contact-based device with a 0°/45° measurement geometry.

The device is turned on. Temporal stability of this device was not able to be conducted due to the design of the 'manual' instrument as well as issues with the software being unable to take consecutive measurements every 30 seconds. The inter-device deviation is +/- 0.10 ΔE*ab on white ceramic so the device will never reach a ∆ E of zero (X-rite, 2021). The device was given 30 minutes to warm up like the Konica Minolta spectrophotometer.

While the device can take and store measurements without software, it requires software to export the data. For all the experiments conducted in this thesis, the device was connected to the software program Color iControl for all measurements taken. This allows for quicker export of the data. A target mask is attached to the instrument shoe as shown in Figure 16. There are three different target masks available: 4mm, 8mm, and 16mm diameter. For all the experiments in this thesis a 4mm target mask

was chosen due to the small size of the samples. The target mask allowed for a 6mm illumination area and a 4mm measurement area.

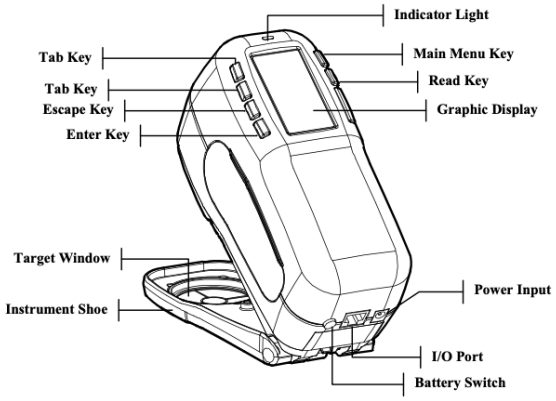

Figure 16: Diagram of Instrument parts (X-rite, 2021)

Measurement conditions for this device include choice of standard illumination (one or more), standard observer, colour space, manual averaging, pass/fail, shade sort, indices, colour difference, opacity, and metamerism (X-rite, 2021). The only measurement conditions that were set were the standard illuminant, standard observer and colour space. The standard illuminant was set to D65, standard observer to 2°. Similar to the Konica Minolta spectrophotometer this information is used so the software can calculate desired colorimetric data (ASTM E1349-06, 2018; Konica Minolta, 2014). However, the colorimetric data in this thesis were calculated from spectral data using MATLAB code.

A two-part calibration occurs after all device settings are chosen: a white calibration followed by a zero calibration. The instrument comes with a reference base. The target mask of the instrument shoe is placed over the white ceramic tile shown in Figure 17.

Figure 17: X-rite Spectrophotometer placed onto provided calibration block. Used the white circle for white calibration and black port for zero calibration.

To take a measurement the device is pressed and held down to the instrument shoe. A measurement takes place and a “Success” is produced on the display screen of the device and will request a zero calibration. The instrument is then placed over the port opening for a zero/black calibration. Device is pressed and held down to the instrument shoe to take the measurement. Display screen will show “Success” when measurement is complete. This indicates the calibration is complete.

After calibration is complete there are two options. A target sample can be measured to which all other samples will be compared to. This producer is used for pass/fail evaluations mostly used for quality control purposes or to check for colour differences. This procedure was not used in this thesis. The second option is to just measure the sample. To measure a sample, a sample was placed over the measurement port of the target mask. The body of the device was pressed into the instrument shoe to trigger a measurement, then is held down as a gas-filled tungsten lamp illuminates the sample and a measurement is collected.

2.2.3 Konica Minolta CS-S100 Spectroradiometer

A tele-spectroradiometer is typically used for the colour measurement of samples emitting light such as display screens (Hunt and Pointer, 2011). It measures irradiance and radiance (Joiner and Luo, 2017). However, the device is capable of measuring reflective objects with the use of external light sources. In this thesis, an illumination cube was used with the spectroradiometer to provide roughly diffuse D65 illumination for each sample. Spectroradiometer is turned on and is allowed to warm up. While the instruction manual states that the device takes a minimum of 20 minutes to warm up, an experiment was conducted to confirm this warm up time (Konica Minolta, 2007).

Figure 18: Calibration Plate

A white Konica Minolta calibration plate was measured every 30 seconds for 4.5 hours (Figure 18). The colour difference (∆Eab) was calculated by computing the difference between the very last CIELAB measurement of the white tile compared every CIELAB measurement between the first and last.

The device reaches temporal stability after 40 minutes shown in Figure 19. However, ∆Eab does not remain at 0 for the duration of 4.5 hours as the instrument has inter-instrument variability of ±0.3 nm for each spectral data or ±2% for luminance measurements (Konica Minolta, 2017).

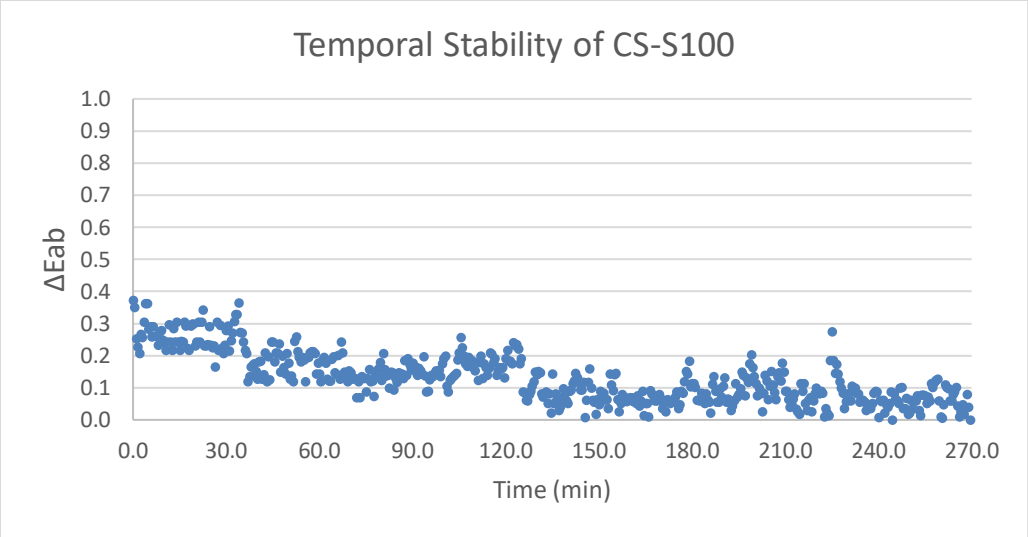

Figure 19: Temporal Stability of Spectroradiometer CS-S100

The device is turned on and allowed 40 minutes so the device to reach temporal stability. The tile was measured for 4.5 hours to certify the device produces reliable measurement for the duration of each experiment conducted in this thesis. The device produces reliable colour measurement for the duration of experiments conducted in this thesis with the largest ΔEab=0.28 after warm up.

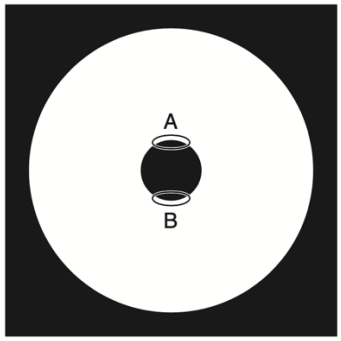

Figure 20: Focusing objective lens for colour measurement

The spectroradiometer is positioned so the focus lens is directed toward the object for measurement. For all the experiments conducted in this experiment, the focus lens was positioned at a 0° measurement angle from the object. Objective lens changes the focal length and is adjusted in order to ensure the sample is focused for colour measurement. Looking through the finder window, the ring is adjusted until A or B are clear in Figure 20.

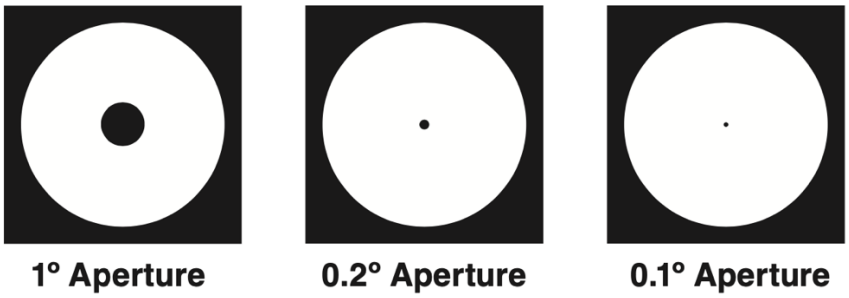

Figure 21: Size of aperture depending on measuring angle

The focal length is dependent on the distance the spectroradiometer is from the sample. Measurement angle selector choices are 1°, 0.2° and 0.1°, which resemble different aperture sizes Figure 21. Depending on how far the sample is away from the end of the objective lens and the measurement angle selected, will change the measurement area (Table 1)

Table 1: Measurement area based on measurement angle and distance (Konica Minolta, 2007)

Measuring Angle	1°	0.2°	0.1°
Measuring Area when measuring distance is 500mm	7.78 mm	1.56 mm	0.78 mm
Measuring Area when measuring distance is 1,000 mm	16.66 mm	3.33 mm	1.66 mm

For all the experiments in this study a 1° measurement angle was used. The distances between the objective lens and sample will change depending on the experiment changing the measurement area.

The device is connected to the PC to use with CS-S10w software. The CS-S10w software is similar to SpectraMatrix and can convert measurements into colorimetric data. This setting was not used, but was used for the collection of radiance data. For each sample, radiance data is collected from 380nm to 780nm at 1nm intervals.

2.2.3.1 Illumination Cube

The VeriVide DigiEye illumination cube can be used with either the spectroradiometer or the Nikon D700 digital camera. It is a totally enclosed, 690mm x 730mm, controlled lighting cabinet. The illumination cube provides either angled (with addition of mirrors) or diffuse illumination by changing the direction of the illumination beams. In all the experiments conducted in this thesis, diffuse mode is used. The controlled lighting cube provides CIE Standardized D65 illumination. The temporal stability of illumination cube was measured with the spectroradiometer. Measurements were taken every 30 seconds for 4 hours of a Konica Minolta white calibration plate. The illumination cube requires 10 minutes to reach temporal stability in which ΔEab maintains a 0.3 over a steady period of time shown in Figure 19 . The device remained consistent over the entire duration of 4 hours.

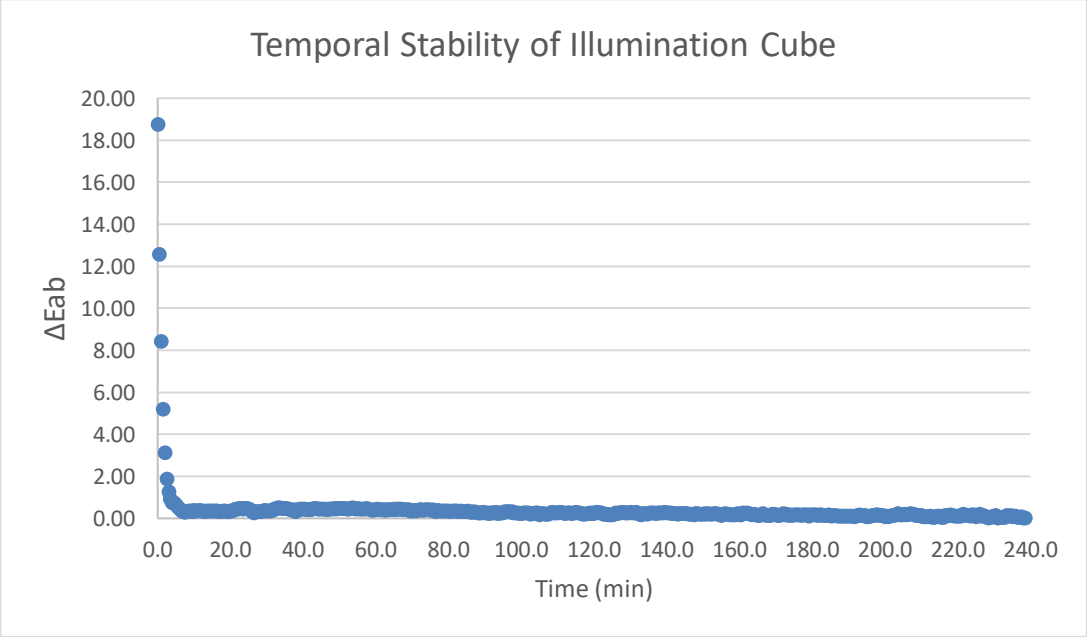

Figure 22: Temporal Stability of Illumination cube over a 40-minute period of time

2.2.4 VeriVide DigiEye System

Digital cameras are non-contact systems which require separate illumination and viewing angle set up (Joiner and Luo, 2017). For dental

photography, a single lens reflex camera with a macro lens is required (Bengel, 2002). The VeriVide DigiEye system is an enclosed illumination cube with standardized CIE D65 illumination with a NikonD700 camera. The Nikon D7000 is a digital single-lens reflex camera that is fitted with a 35mm 1:2D lens. Images of teeth are taken up close requiring a macro lens, which the Nikon does with the 35mm lens (Bengle, 2002). For all experiments conducted in this thesis, the camera was set up to take images with autofocus. The digital camera was set to have an aperture of 5 and a shutter speed of 1/8. DigiEye illumination cube was turned on and allowed 10 minutes to warm up as suggested in section 2.2.3.1. Nikon D7000 is turned on and is connected to the DigiEye software.

Figure 23: White Calibration Tile

Prior to measuring the samples, the DigiEye system is calibrated. The Nikon D7000 is switched to manual focus and a white tile provided by VeriVide is measured (Figure 23). This portion of the calibration allows for white balance to be adjusted in order to remove unwanted colour castes as well as to estimate the colour of the illumination and correct for it (Westland *et al.,* 2012). It also helps to ensure the repeatability of captured images (Zhang *et al.,* 2017). Digital camera is switched to autofocus for the measurement of a colour calibration chart, a chart with references colours with known XYZ values. Custom made colour calibration chart

made by DigiEye was used for the digital camera characterisation. The DigiTizer Chart V3.91 contains 237 patches with a central white patch (Figure 24). Proper colour adjustments are essential to acquiring more precise colour measurements (Wee *et al.*, 2006; ASTM E2466-13, 2013).

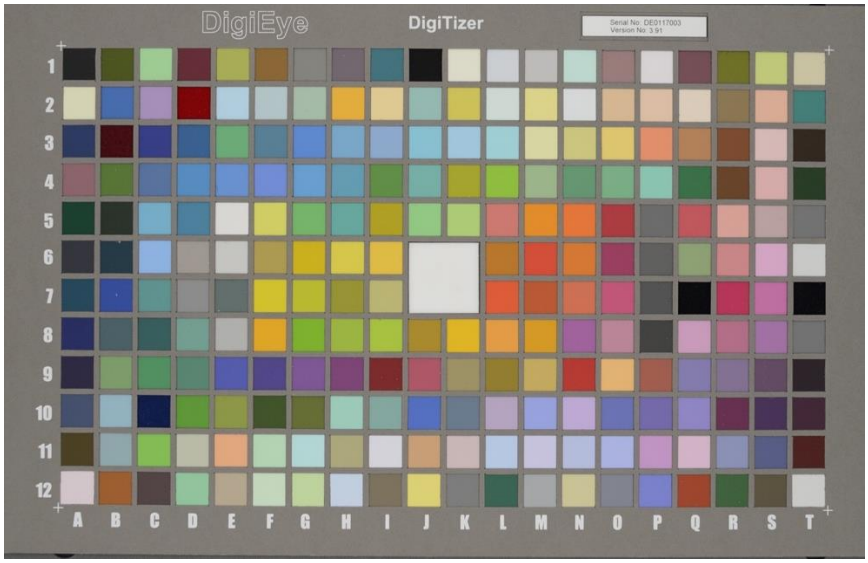

Figure 24: Colour Calibration Chart V3.91

Light reflected from the sample is collected by photosensors of the camera and converted into red, green and blue (RGB) values (Martínez-Verdú *et al.*, 2010). The CIE developed a standardised method for converting RGB to XYZ using the following equation (Hunt and Pointer, 2011):

$$\boldsymbol{X = 0.49R + 0.31G + 0.20B}$$

$$\boldsymbol{Y = 0.17697R + 0.81240G + 0.01063B}$$

$$\boldsymbol{Z = 0.00R + 0.01G + 0.99B} \qquad \textbf{Equation 2-1}$$

However, camera characterisation and colour calibration charts will influence this equation. In order to convert the RGB values into device independent XYZ values, a camera characterisation model is constructed based on a linear transform (Mohan *et al.*, 2008; Joiner and Luo, 2017). Characterisation is done by imaging a colour chart with a set of known tristimulus values (Westland *et al.,* 2012; Martínez-Verdú *et al.*, 2010). A complete characterisation includes spatial characterisation which involves

the application of a linear correction algorithm to compensate for spatial non-uniformity of the camera sensor as well as spectral characterisation which involves obtaining RGB colour matching functions of the camera (Martínez-Verdú *et al.*, 2010). The colour characterisation is a function provided by the DigiEye software and a numerically optimised equation for CIEXYZ from RGB values are generated for every calibration.

After calibration, the sample is placed inside the DigiEye illumination cube. Placement varied depending on experiment. The Nikon D700 was placed on auto- focus for measurements of samples. The benefit of autofocus is easier operation. An image was captured of the sample. Colour measurements were taken of the sample and the software produces a set of CIEXYZ values produced from the averaging of RGB values of every pixel in that measured area that are converted into one set of XYZ values. Three measurements were taken for each tooth and average to produce one set of XYZ values for each tooth.

2.2.5 Vita EasyShade Advanced 4.0

The Vita EasyShade Advanced 4.0 is trade-marked spectrophotometer specifically designed for tooth shade determination of natural teeth and ceramic restorations (VITA, 2020b; Rauf, 2020). It is a contact-based system with an internal D65 light source. The Vita EasyShade device and Vita EasyShade helper software allow for the quick determination of tooth shade based on their database containing shade guide standards of VITA classical shade guide (A1-D4), the VITA Linear 3D-Master shade guides, Vita Toothguide 3D-Master, and Bleached Shade guide (VITA, 2020a).

EasyShade Advanced 4.0 is turned on by pressing and holding any button on the handpiece for at least two seconds. Ensure the hand-piece is warmed up. Based on the instruction manual for the device (VITA Zahnfabrik, 2013), there is no designated amount of time for the device to warm up. Usually, measurements are taken of a white tile for an allotted amount of time in order to determine the time needed to reach temporal

stability. This device is connected to software called VITA EasyShade Helped and is not equipped to measure opaque samples. It is unknown how long it will take for the device to reach temporal stability. Since the EasyShade is a spectrophotometer, it was allotted 5 minutes for the device to completely warm up like the Konica Minolta CM-2600d. Clinical procedure is to turn on the device, calibrate it and take a few test measurements of known standards prior to taking measurements of a patient (Hogan, 2021).

In between patients, it is important to clean and disinfect the instrument to prevent patient cross-contamination. After warm up, an infection control shield is applied to the probe of the hand-piece. An infection control shield is a disposable plastic shield to prevent cross-contamination. The tip of the handpiece is inserted into the shield and is stretched uniformly and flat over the face of the probe pulled carefully over the hook on the hand piece in order to keep the shield secure (VITA Zahnfabrik. 2013). Although the device was only used on ceramic samples in the experiments in this thesis and cross-contamination was not a problem, an infection control shield was still applied to replicate the basic procedure which occurs in dental offices. The control shield was not replaced in between ceramic samples as there was no concern about cross-contamination.

Figure 25: Demonstration of the Vita EasyShade Advance 4.0 calibration

With the device warmed up and an infection shield in place, the unit is calibrated. The device can be calibrated manually or automatically. The probe of the hand-held unit is placed flush with and perpendicular to the calibration block which is constructed into the base (Figure 25). When calibration block is compressed activating automatic calibration of the device when connected to a power source. Otherwise, the measurement switch must be pressed to manually calibrate den all experiments, the device was automatically calibrated (VITA Zahnfabrik, 2013).

There are four modes of operation: single tooth mode, measurement of tooth areas, averaged measurement, and measurement on a ceramic restoration. Basic shade measurement is designed to only measure the central area on the tooth and can determine the general tooth shade of natural teeth. Measurement of tooth areas enables the device to determine the basic tooth shade of three areas of the tooth: the cervical, central and incisal (VITA Zahnfabrik, 2013). Average measurement calculates an average basic tooth shade form several measurements. Measurement of ceramic restorations is used for comparing the shade of a ceramic restoration with a tooth shade stored in the EasyShade Advance 4.0 Helper software against their database of shades, for comparing the shade of a ceramic restoration with a previously measured tooth shade, or determination of the shade of a ceramic restoration. For the experiments using the EasyShade in this thesis, single tooth measurement mode was selected.

In a clinical setting, the patient would sit in the chair with their head leaned back. Dentist would locate colour gradient of tooth, designate the area to measure, place the probe tip flush against the tooth surface and then collect a measurement. In the case of the experiments conducted in this experiment, the probe was held flush with the sample. While holding the tip steady against the tooth, shade guide or sample, measurement button is pressed.

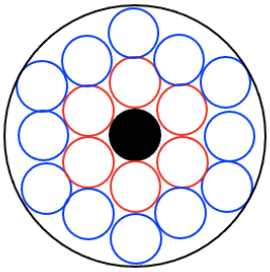

Figure 26: Diagram of the probe of EasyShade with the detector (black), inner array of LED lights (red), and the outer array of LED lights (blue)

The probe has two rings of LED lights, WITH 6500K temperature, surrounding the detector. Each LED light and the detector have a diameter of 1 mm. The entire probe has a 5 mm diameter. Two consecutive white high-powered LED lights with a colour temperature of 6500K will flash (Rauf, 2020). The two flashes capture different measurements, an inner and an outer measurement, which correlates to the inner and outer circle of LED lights (Figure 26). This device does not follow any CIE standardized measurement geometry. The inner section of lights flashes first, and the outer measurement occurs second in rapid succession to the first measurement. Body reflected light is measured by the eleven spectral channel sensors located in the centre of the probe. The sensor is equipped with spectral interference filters capable of converting light into frequency (Rauf, 2020). Two rapid “beeps” can be heard indicating the completion of the measurement (VITA Zahnfabrik, 2013).

Vita EasyShade Helper software provides details of measurements taken such as method (inner and outer measurement), spectral data from 400nm to 700nm at 10nm intervals, CIELAB colorimetric values, CIE LhC, CIE XYZ values, shade match prediction for Vita Classic, shade match prediction for Linear 3D-Master, shade match prediction for Toothguide 3D-Master, and shade match prediction for Bleached guide (VITA Zahnfabrik, 2013). Software created by Vita was designed to utilise artificial intelligence to classify spectral data and assigns it a dental shade

to the recorded special data (Rauf, 2020). Spectral data collected from the device was used for calculations in this thesis. Each sample was measured three times producing a total of six measurements; three inner measurements and three outer measurements. Those measurements were averaged to produce a single set of spectral data for each sample to be used for the conversion to colorimetric data.

2.3 Open Data

Open data refer to freely available data for use by all researchers (i.e. international research communities and sometimes general public), whether or not it has been published or not (Pasquetto *et al.*, 2016; Hughes, 2017). While life sciences and earth sciences have shared dated for years, public access to other fields of data is rapidly moving towards becoming the norm (Piwowar and Vision, 2013; He and Nahar, 2016). Open data refers to 'research data' versus government statistics and industry records which includes data about culture, scientific, financial, statistical, weather, environmental, transport, medical (Pasquetto *et al.*, 2016). Sharing data or reusing data means using research data that for analysis for the purpose other than it was originally intended facilitating science (National Library of Medicine, 2021; Piwowar and Vision, 2013). Some benefits of re-analysing data in an innovative or different way include gaining multiple perspectives, exploration of topics not envisioned by initial research, identifying errors, validation research results, deterring inaccurate reporting in clinical trials, increasing efficient use of funding, allows for creation of new datasets by combining data from multiple sources and patient population resources by avoiding duplicate data collections (Piwowar and Vision, 2013; Pasquetto *et al.*, 2016; Mello *et al.*, 2018; Hughes, 2017; He and Nahar, 2016).

Raw data collected by fellow colleagues at the University of Leeds doing research for Colgate Palmolive was available to be re-used. Already published data by QianQian Pan, Stephen Westland and Roger Ellwood

(2018) was used in chapter 3 and chapter 5. Unpublished and unanalysed raw data collected by the same group based on perceptual yellowness was available for analysis in the same chapters. Accessing raw data which has been unpublished is considered at asset as it reduces the use of resources such as funding as well as the data does not go to 'waste' (Hughes, 2017). Re-using data allows for the exploration of related or new hypothesis (Piwowar and Vision, 2013). The raw data being re-used in this thesis allows for the exploration of related hypothesis, building upon previous findings, and accelerate new information regarding tooth colour (Mello *et al.*, 2018; Pasquetto *et al.*, 2016). From a business perspective, re-using the data from pervious Colgate research to build upon is cost effective having to reduce resources. An additional benefit is not having to duplicate procedures. One of the experiments conducted by Pan and colleagues (2018), data from 500 participants was collected from 5 different countries. Another experiment in the same study collected data from 80 participants from China and the UK. Trying to replicate collecting this amount of data would waste time and money when the raw data is readily available.

Concerns about open data mainly focus on data protection. It is crucial legally that no potential harm to research participants can occur and ensuring their privacy is adequately protected especially sensitive information such as sexual orientation or substance abuse (Mello *et al.*, 2018; Piwowar and Vision, 2013). Another major issue with open data is that there are legal concerns of ownership of data sets (Pasquetto *et al.*, 2016). The raw data re-used in this thesis was anonymous when collected. Data was collected by Colgate researchers working with the University of Leeds and is being re-used by a Colgate researcher at the University of Leeds. There are no problems regards legal ownership of the data as there is already a legal contract between Colgate and the University of Leeds in regards to the data that is collected during their collaboration. Multiple people handle research in corporations as do many research groups in

Universities across the world and should be acceptable for this study. Data sharing is encouraged in countries such as the USA, UK, and Australia. It has been found that papers that reused data belonged to authors within their own research groups in various disciplines (i.e. biological science, agricultural, genetics, medicine, etc.) (He and Nahar, 2016). A concern for reusing data includes the reliability of the data. Raw data used in this thesis was collected by a colleague in the Colgate-University of Leeds research group and documents detailing data collection were provided. According to He and Nahar (2016), data collected within the same research group is more credible as detailed records are provided to insure reliability of data. The data being re-analysed in this thesis was published, therefore has been peer reviewed and accepted. Unpublished raw data used in this thesis was collected at the same time as the data that was published by Pan and colleagues (2018).

2.4 Psychophysical Methods

Scaling Analysis is a psychophysical technique that describes the relationship between the physical magnitude of stimuli and the perceptual responses to the same stimuli (Kingdom and Prins, 2010). In dentistry, paired comparison and rank order are the two visual methodologies typically used for the quantification of ‘whiteness’ and ‘yellowness’ attributes of teeth (Hirschen, 2010; Luo *et al.*, 2009; Pan *et al.*, 2018; Sullivan *et al.*, 2019; Pan and Westland, 2018; Pérez *et al.*, 2016). These two methods produce perceptual scales which are descriptions of stimulus appearance and have procedures that have no correct or incorrect answer (Kingdom and Prins, 2010).

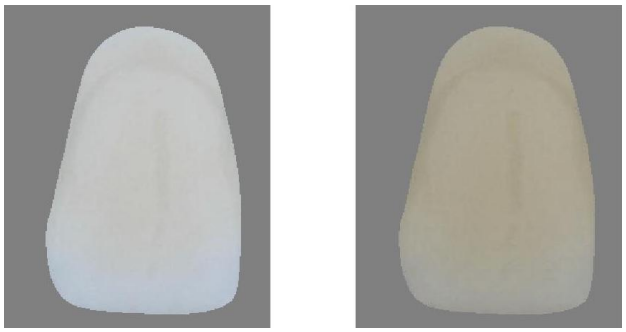

Figure 27: Paired Comparison of two tooth images

Paired comparison is a forced-choice method in which participants are not allowed a 'non-response' answer and are required to make a choice between the provided stimuli (Kingdom and Prins, 2010; Howitt and Cramer, 2014). Two stimuli are presented at one time and volunteers are asked to pick one based on a question such as 'which image or tooth is whiter' (Tsukida and Gupta, 2011) (Figure 27). Ideally comparisons are made for all possible pairs (Tsukida and Gupta, 2011). If the procedure is repeated for all possible pairs of samples, then the samples can be rank-ordered (Kingdom and Prins, 2010).

Figure 28: Example of a set of stimuli presented to a participant to rank based on a specific appearance attribute

Rank order, also known as multi-stimulus scaling, is a method in which an observer is presented with an entire set of stimuli together. The observer is asked to arrange the set of samples in a ranking based on some appearance attribute such as whiteness or yellowness (Figure 28) (Kingdom and Prins, 2010). Like paired comparison, this method is also 'forced-choice'. Both paired comparison and rank order provide ordinal data, which provides information regarding the order of preference or perceptual attribute, such as which one image is whiter, but it does not reveal the magnitude of how much whiter (Luo *et al.*, 2009; Howitt and Cramer, 2014).

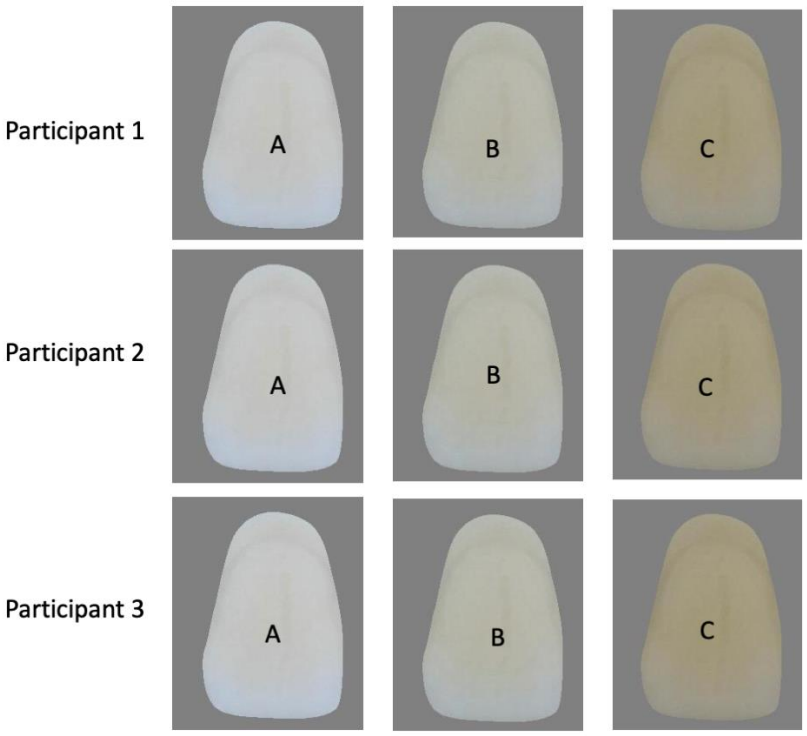

Figure 29: Example of three participants arranging samples based on whiteness

Figure 29 shows three participants' ranking three tooth samples based on whiteness with the whitest sample on the left to the least white on the right. If the differences between samples A, B, and C are large enough every participant will arrange the samples in the same order. In this case, only information on the order of whiteness is provided: A is whiter than B and B is whiter than C. This is ordinal data. The magnitude of how much whiter A is than B versus how much whiter B is than C is unknown. Interval scale data is needed, which is a numerical scale that corresponds to relative perceptual differences among the stimuli (Kingdom and Prins, 2010; Luo *et al.,* 2009).

To generate interval data instead of ordinal data, the colour differences between stimuli must be small enough ensuring the response to any pair or ranking are not always the same.

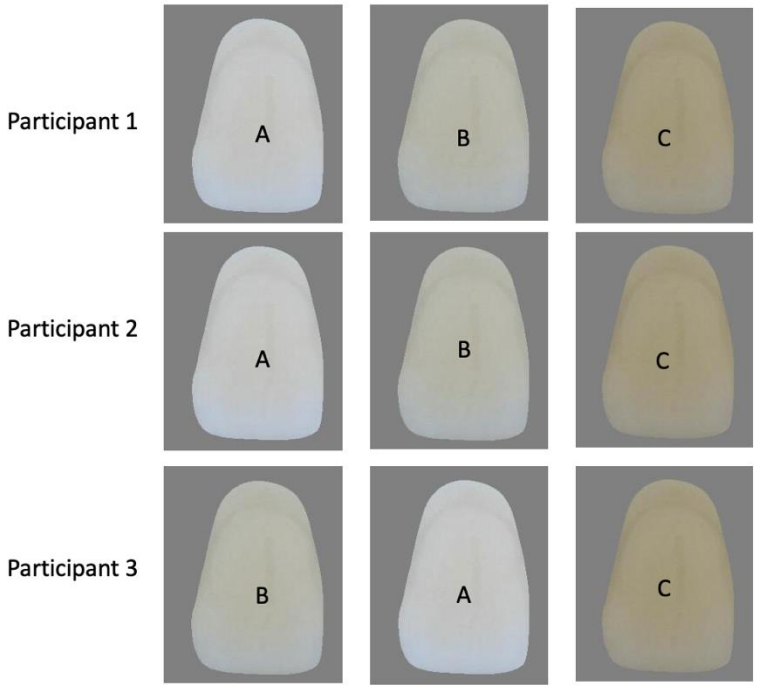

Figure 30: Example of three participants' varying responses to whiteness

For example in Figure 30, three participants show varying responses to the order of whiteness. The variation in responses provide meaningful information regarding perceptual differences. Participant three arranged samples A and B differently than everyone else. This indicates that sample A and B are perceptually very similar. All three participants responded that sample C is the least white. The variation in responses provides information indicating that the perceptual difference between samples C and A or B is greater than the perceptual differences between A and B. The limitation of interval data is that it does not capture the magnitude of these differences; thus, it cannot be stated that A is 1x whiter than B, and A is 5x whiter than C (Kingdom and Prins, 2010). However, the differences create proportion data. A stimuli is chosen a proportion of times over another stimuli allowing for an estimate of perceptual differences between stimuli, which are converted into a standard unit of measurement for communication called Z-scores (Kingdom and Prins, 2010; Howitt and Cramer, 2014). The calculation method for Z-scores is explained in section 2.4.1.

2.4.1 Scaling Analysis (Z Score)

A Z-score (Z) is a statistical method which provides a standard unit of measurement for analysis. The number of standard deviations from the mean provides a numerical expression of the relative differences amongst stimuli (Howitt and Cramer, 2014; Luo *et al.*, 2009; Kingdom and Prins, 2010). The advantage of this method is that the number of standard deviations is equally applicable to a variety of variables such as anxiety, time, weight, or whiteness and yellowness. This allows for different units of measurement to be easily compared with each other (Howitt and Cramer, 2014).

Figure 31: Nine tooth samples varying in whiteness

For example, a set 6 participants are asked to rank 9 tooth samples based on whiteness as shown in Figure 31. The most 'white' sample perceptibly is given a 1 and the least white samples is given a 9. The basic idea of an interval scale is that the number valued between each sample are equal in size, even though the actual number is arbitrary (Howitt and Cramer, 2014). In this case, the location in which the sample is ranked is equal in size. Table 2 displays an example of how each tooth was ranked by each observer. Observer 1 perceived tooth E as the 'whitest' sample out of the set, with F as the least white.

Table 2: Example of the rank order of four tooth samples by five different participants

Tooth Sample	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5	Observer 6
A	5	8	6	5	5	5
B	6	9	8	7	8	9
C	3	2	3	8	2	3
D	8	7	9	4	9	6
E	1	4	5	2	7	4
F	9	5	2	6	4	8
G	2	1	1	9	1	2
H	7	6	7	3	6	7
I	4	3	4	1	3	1

To produce a Z score, the average placement of the stimuli by all the observers based on their perception of 'whiteness' is calculated. Each averaged rank order is converted into proportion ratios by the equation (N-K)/(N-1), turning them into a percentage value. N represents the total number of stimuli in front of each participant at a time. In the example above, N would be equal to 9. K is the average perceptual ranking of the stimuli. This portion of the calculation represents the comparison of each sample with one another in that specific set (N-1), which is equal to 8 in the example above. The perceptual ratios were converted to generate a corresponding probability value by calculating the inverse of the cumulative standardised normal distribution, using the normsinv function in Microsoft Excel.

Table 3: Calculated Z scores for each tooth based on the responses of perceptual whiteness from 6 observers

Tooth Sample	Mean Ranking	Percent Proportion (N-K/N-1)	Z-Score
A	4.33	0.58	0.21
B	1.50	0.94	1.53
C	6.17	0.35	-0.37
D	2.33	0.83	0.97
E	4.50	0.56	0.16
F	5.83	0.40	-0.26
G	8.17	0.10	-1.26
H	3.83	0.65	0.37
I	8.33	0.08	-1.38

The standard normal distribution function produces a numerical range in which the mean value becomes 0 (Howitt and Cramer, 2010; Kingdom and Prins, 2010; Thurstone, 1927). Normalizing the distribution allows for the data to be understand easier. A positive value away from 0 indicates the standard deviation is above the mean describing a whiter (depending on perceptual assessment) sample and a negative value from 0 indicates the standard deviation is below the mean less white sample. Based on Table 3, perceptually the sample that is most white is tooth I, whereas the least white sample is B. This statistical method allows for perceptual response to be calculated and compared to physical colour measurement data in a quantitative way. The Z-scores correspond to the relative perceptual differences between stimuli (Luo *et al.*, 2009).

2.5 Data Preparation

Each instrument collects data that needs to be translated into a useful form of colorimetric data for dental professionals (Chu *et al.*, 2010). Spectrophotometers do not measure perfect reflectance, but rather a reflectance factor (i.e. values between 0 and 1), a ratio of reflected light of reference white standard compared to a prefect white which reflects 100% at all wavelengths (ASTM E179-17, 2017; Luo, 2011; Wee *et al.*, 2006). The software calculates the reflectance factor using the following equation:

$$\boldsymbol{R_{(\lambda)} = \frac{P_{i(\lambda)}}{P_{w(\lambda)}}} \qquad \textbf{\textit{Equation 2-2}}$$

where R(λ) represents the adjusted reflectance, Pi(λ) is the reflectance of sample, and Pw(λ) is the reflectance of a reference standard white (i.e. measured during the calibration process).

However, spectroradiometers measure spectral radiance instead of reflectance. Spectral radiance data must be converted into reflectance

factor data in order to calculate useful colorimetric values. The radiance data is converted into reflectance factors using the following equation:

$$R(\lambda) = \frac{S_{i(\lambda)}}{S_{w(\lambda)}} * P_w \qquad \textit{Equation 2-3}$$

In which R(λ) describes the converted reflectance data, Si(λ) is the radiance data collected of sample, Sw(λ) is the spectral radiance data of a reference standard white, and Pw is reflectance data of the same reference white standard measured with a spectrophotometer. A Konica Minolta Calibration plate, shown in Figure 32, was the white standard measured by the spectroradiometer and spectrophotometer for the calculation. The averaged specular excluded spectrophotometer measurements of the white plate were used, as the geometry of the spectroradiometer was always positioned at an exact 0° angle from the sample preventing the specular component to be included.

Figure 32: Standard white calibration plate

The spectral data range needed to be the same for the spectroradiometer and spectrophotometer. The spectroradiometer collects radiance data from 380 nm to 780 nm at 1nm intervals and the spectrophotometer collects reflectance data from 360 nm to 740 nm at 10 nm intervals. Only data from 380 nm to 740 nm are used for the calculation. Since the spectroradiometer collects data at 1nm intervals, the spectrophotometer data must be interpolated from 10nm interval to a 1 nm interval.

Interpolation is a numerical method in which constructs new data points within the range of a set of known data points (Westland *et al.*, 2012).

2.6 Conversion of spectral data to colorimetric data

Each averaged measurement collected by the different instruments were converted into three different colour spaces: CIEXYZ, chromaticity coordinates, and CIELAB.

2.6.1 Tristimulus Values (XYZ)

Measurements from the DigiEye system produces XYZ values for each sample and therefore these do not need to be manually calculated. However, the spectral data collected from the spectroradiometer, spectrophotometers, and EasyShade must be converted into XYZ values. The reflectance data was converted into CIEXYZ values using equation 1-1 in section 1.5.1. In this experiment the standard observer is always 2° and standard illuminant is D65. Weighing tables can found in Wyszecki and Stiles (1967) book Color Science or from the American Society for Testing and Materials in E308-01 (ASTM, 2001).

XYZ values were computed using MATLAB code r2xyz (Appendix A) from a downloaded MATLAB colour toolbox (Westland *et al.*, 2012). The standard observer and standard illuminate are inserted into the code as ‘d65_31’, which allows for the software to access its internal database of ASTM E308-01 (2001) weighting tables for a 2° observer for D65 illumination of a standard white for the inserted spectral range of 360nm to 740nm at a 10 nm interval. The measured spectral data is also inserted into the code, and the software exports the calculates XYZ values for each sample. This code only works for spectral measurements collected at 10nm intervals.

The spectroradiometer collects data at a 1nm interval. Modification to the r2xyz code and weighing table for ‘d65_31’ were required in order to

calculate XYZ values based off a 1nm interval instead of a 10nm interval. The ‘d65_31’ weighing table was interpolated in order to produce weighted data at 1nm intervals instead of 10nm intervals. The modification was named r2xyz_mod.m and is shown in Appendix B.

2.6.2 Chromaticity Coordinates (xyz)

The equations 1-2 in section 1.5.2 were used to calculate the chromaticity coordinates from CIEXYZ.

2.6.3 CIELAB (LAB)

All CIEXYZ data were also converted into CIELAB values. The CIELAB values can be calculated from tristimulus values using equation 1-3 in section 1.5.3. All L*, a*, and b* values were calculated in MATLAB using code xyz2lab.mat (Appendix C). The XYZ values for each sample were inserted into the code with a 2° observer for D65 illumination which is coded as ‘d65_31’.

2.7 Colour Difference

Colour difference (∆E) is a quantitative representation describing difference between colour measurements (Khashayar *et al.*, 2014). The aim of a colour difference formula is to produce values which relate to visual perception differences (Lee and Powers, 2005). The higher the ∆E value, the larger difference in colour to the human eye (Khashayar *et al.*, 2014). In dentistry, colour difference is used for the evaluation of colour replication of dental restorations, colour instability of dental polymers, colour perceptibility and acceptability, translucency parameters, colour changes by processing dental materials, and evaluation of colour specifications of prosthetic materials (Wee and Lindsey, 2006; Khashayar *et al.,* 2014; Johnston, 2009; Oliveira *et al.,* 2015; Zenthöfer *et al.,* 2014; Mazu et al, 2020). Three commonly used colour difference equations are: CIELAB (∆Eab), CMC, and CIEDE 2000 (∆E00). In dental research, ∆Eab is the most commonly used for the evaluation of tooth colour (Lee and

Powers, 2005; Baltzer and Kaufmann-Jinoian, 2004; Da Silva *et al.,* 2008; Kim *et al.*,2013; Lim *et al.,* 2010). Colour difference CIELAB (∆Eab) is calculated by evaluating the differences between two sets of LAB data shown in Equation 0-5; ∆L is the difference between two CIELAB L* values, ∆a* is the difference between two CIELAB a* values, and b* is the difference between two CIELAB b* values. To calculate colour difference (∆Eab), MATLAB code cielabde.m was used shown in Appendix D.

$$\Delta \boldsymbol{E^*ab} = \sqrt{\Delta \boldsymbol{L^{*2}} + \Delta \boldsymbol{a^{*2}} + \Delta \boldsymbol{b^{*2}}} \qquad \textit{Equation 2-4}$$

Even though colour difference ∆Eab is more commonly used and easier to calculate, it is suggested that colour difference ∆E00 is more useful in dentistry both in research and clinically as it reflects the colour differences perceived by the human eye better (Gómez-Polo et al., 2015). Luo et al. (2001), developed the colour difference formula ∆E00 for the evaluation of small colour differences such as changes in tooth colour. Colour difference (∆E00) has been used in evaluating the colour stability of colour resins and evaluating the influence of dental occlusions on tooth colour determinations (Mazur et al., 2020; Oliveira et al., 2015). The ∆E00 formula uses lightness (L), chroma (C*), and hue (h*) colorimetric values instead of CIELAB. Chroma is the intensity of a colour (saturation) and hue is shade (i.e. red, green, blue, etc). Chroma and hue can be calculated from CIELAB values using equation 2-5.

$$\boldsymbol{C^*_{ab}} = \sqrt{\boldsymbol{a^{*2}} + \boldsymbol{b^{*2}}}$$

$$\boldsymbol{h_{ab}} = \boldsymbol{tan^{-1}}\left(\frac{\boldsymbol{b^*}}{\boldsymbol{a^*}}\right)\left(\frac{\boldsymbol{180}}{\boldsymbol{\pi}}\right) \qquad \textbf{\textit{Equation 2-5}}$$

To calculate ∆E00, equation 0-8 was used (Luo *et al.*, 2001).

$$\Delta \boldsymbol{E00} = \sqrt{\left(\frac{\Delta \boldsymbol{L'}}{\boldsymbol{k_L S_L}}\right)^2 + \left(\frac{\Delta \boldsymbol{C'}}{\boldsymbol{k_C S_C}}\right)^2 + \left(\frac{\Delta \boldsymbol{h'}}{\boldsymbol{k_h S_h}}\right)^2 + \boldsymbol{R_t}\left(\frac{\Delta \boldsymbol{C'}}{\boldsymbol{k_C S_C}}\right)\left(\frac{\Delta \boldsymbol{C'}}{\boldsymbol{k_C S_C}}\right)} \qquad \textbf{\textit{Equation 2-6}}$$

Where

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

And

$$S_C = 1 + 0.045\bar{C}'$$

And

$$S_H = 1 + 0.15\bar{C}'T$$

Where

$$T = 1 - 0.17\cos(\bar{h}' - 30°) + 0.24\cos(2\bar{h}') + 0.32\cos(3\bar{h}' + 6°) - 0.20\cos(4\bar{h}' - 63°)$$

And

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

Where

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

And

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

The $\bar{L}'$, $\bar{C}'$, $\bar{h}'$ are the means of the L', c' and h' values for a pair of samples (Luo et al., 2001). Weighing functions are included in equation for lightness (SL), chroma (SC), and hue (Sh) that account for differences between chroma and hue. The RT is the rotation term which accounts for differences in chroma and hue for saturated colours in the blue region. Parametric factors (kL,kC,kh) correct for influences due to experimental viewing conditions on perceived colour differences (Melgosa et al., 2013). The MATLAB code ciede00.mat from colour toolbox was used to calculate all ΔE00 values shown in Appendix E (Westland et al., 2012). The code calculates hue and chroma within the calculation. Two different sets of CIELAB values for comparison are input into the code as well as 1 for SL, SC, and Sh, so the calculation uses the default values for the weights.

2.8 Indices

Whiteness indices WIO, WIC and WID are used in thesis as well as yellowness indices YIE313 and YID259. Another yellowness index, not

mentioned in section 1.7, is YIO which is developed in chapter 3 and is used for analysis in chapter 5. The YIO index :

$$\boldsymbol{YIO = -Y - 851.716(x_n - x) - 436.962(y_n - y)} \quad \textbf{\textit{Equation 2-7}}$$

Each index requires specific colorimetric data. Chromaticity coordinates are required from both the light source and sample for the indices WIO, WIC and YIO. YIE313 and YID1925 require CIEXYZ values, while WID requires CIELAB.

For indices WIC, WIO, and YIO the chromaticity values for both the standardised light source and sample are required. According to ASTM E2545-07 (2007) standardised CIE D50 illumination is used for tooth measurement. However, D65 is widely used in dental research (Wee *et al.*, 2006; Pan *et al.* 2018; VITA, 2020b). Most whiteness indices and yellowness indices were created based on D65 illumination which corresponds to the sum of the spectrum of daylight in Western/Northern Europe (Hirschen, 2010; Joiner *et al.*, 2008a). D65 illumination is used as the standard illumination for every experiment in this thesis and for the calculations of all the indices. The x_n and y_n are the chromaticity coordinates for the CIE standard illuminant D65, which are x_n= 0.3127 and y_n=0.3290. These values were calculated from D65 XYZ values using Eqn. 1-1 (Westland *et al.,* 2012). While Y, x, and y are the chromaticity coordinates of the sample. Equations 1-4, 1-5, 2-7 display indices WIC, WIO and YIO. For yellowness indices YIE313 and YID1925, CIE XYZ colorimetric data from the sample are used for calculations (Eqn.1-7 and Eqn. 1-8). CIELAB values are only used for the calculation of the whiteness index WI_D (Eqn. 1-6). Respective colorimetric values were inputted into each equation index value was outputted for comparison. A higher WIC, WIO, and WID value indicates higher whiteness or least yellow of the object (Joiner and Luo, 2017; Luo *et al.,* 2009). A higher YIE313, YID1925, YIO indicates a higher yellowness or least white.

2.9 Statistical Analysis

Statistical methods are used in order to quantify the strength of the relationship between visual perception of colour differences and computed colour differences (indices) from instrumental measurements relationship (Kirchner and Dekker, 2011). The most commonly used statistical methods in dentistry to compare perceptual colour differences and indices are Coefficient of determination (r^2) and percent wrong decision (%WD) (Luo *et al.*, 2009; Pan *et al.*, 2018; Klotz *et al.,* 2018). STRESS is a statistical analysis that has only been used in one dentistry study, which is a publication based on the yellowness index developed in chapter 3 (Sullivan *et al.*, 2019).

2.9.1 Coefficient of determination (r^2)

Coefficient of determination (r^2) is an index of the variance between two variables providing a portion results that is between values 0 and 1. It is a metric which is widely used to measured goodness of fit and has been used in dental research (Kirchner and Dekker, 2011).

To calculate r^2, Eqn. 2-8 is used:

$$\boldsymbol{r^2} = \left(\frac{\sum(\boldsymbol{V}-\bar{\boldsymbol{V}})(\boldsymbol{I}-\bar{\boldsymbol{I}})}{\sqrt{\boldsymbol{\Sigma}(\boldsymbol{V}-\bar{\boldsymbol{V}})^2}\sqrt{\boldsymbol{\Sigma}(\boldsymbol{I}-\bar{\boldsymbol{I}})^2}}\right)^2 \qquad \textit{Equation 2-8}$$

In which V is the visual scale values (Z scores), $\bar{V}$ is the mean score of the visual scale values, I is the computed index (i.e. WIO, WIC), and $\bar{I}$ is the mean score of the computed index. A value close to 1 is regarded as the variables being correlated and a value closer to zero indicating no relationship between the variables (Howitt and Cramer, 2014).

2.9.2 Percent Wrong Decision (%WD)

Percent wrong decision is a method where each stimulus in a data set is compared to every other stimuli in the same data set. The percentage is calculated based off the number of times an observer or metric would disagree about which stimuli out of the presented pairwise comparison is

'whitest' or 'yellowest' compared to the average visual decision from a group of observers (Pérez *et al.*, 2016; McDonald, 1997; Pan *et al.*, 2018). Using the same 9 tooth samples in Figure 31 as an example, Table 4 shows the average visual decision (Z score) and calculated whiteness index (WIO).

Table 4: The Z score value and WIO index for set of 9 stimuli

Tooth Sample	Z score	WIO
A	0.21	67.09
B	1.53	71.75
C	-0.37	64.36
D	0.97	76.10
E	0.16	68.68
F	-0.26	66.06
G	-1.26	58.20
H	0.37	70.41
I	-1.38	62.82

In Table 4, tooth sample A with a z score of 0.21 is considered visually less white than sample B with a z score of 1.53. Since B has a larger z-score it means that the sample is white. The WIO value reflects the same with A having a WIO value of 67.09 and B with a WIO value of 71.75. This is considered a correct answer. Tooth sample A is visually whiter than tooth sample E, however the WIO value states that sample E should be whiter than tooth sample A. This is considered a wrong answer. There are a total of 36 comparisons are evaluated. Out of this example set, there are 3 wrong decisions made out of 36 comparisons, Figure 33.

A vs. B	A vs. H	B vs. G	C vs. G	D vs. H	F vs. G
A vs. C	A vs. I	B vs. H	C vs. H	D vs. I	F vs. H
A vs. D	B vs. C	B vs. I	C vs. I	E vs. F	F vs. I
A vs. E	B vs. D	C vs. D	D vs. E	E vs. G	G vs. H
A vs. F	B vs. E	C vs. E	D vs. F	E vs. H	G vs. I
A vs. G	B vs. F	C vs. F	D vs. G	E vs. I	H vs. I

Figure 33: All possible comparisons. Red highlighted comparisons are the ones in which WIO made a wrong decision compared to visual perception (Z score)

This ratio is multiplied by 100 to get a percentage. In this example set, there is 8.33% wrong decisions. When evaluating a yellowness index (i.e YIE313, YID1925, YIO and b*) compared to the perceptual whiteness, the number of correct comparisons is inverted. An inversion is required as %WD is calculated comparing the arrangement of the highest value based on perceptual whiteness (Z score) and the arrangement of the highest value of the index. For a yellowness index, a larger index value indicates a more ‘yellow’ sample or the least ‘white’ sample. Therefore the number of correct comparisons the yellowness index makes compared to the Z score is actually the number of incorrect comparisons. For example, if the number of correct responses for a yellowness index (YIO) compared to whiteness (z score) is 1218 out of 1326 comparisons. The actual the number of correct comparisons is 1326 minus 1218, which is 108. The percent wrong decision is 8.14%. This applies to when a whiteness index is compared to perceptual yellowness (Z score).

When comparing indices to Z scores, the fewer wrong decisions the more accurate the indices is to visual perception. Percent wrong decisions is a typical statistical method used in dental research (Pan *et al.*, 2018; Pérez *et al.*, 2016; Luo *et al.*, 2009). In this thesis, the value of each tooth sample from each of the indices (WIO, WIC, WID, YIE313, YID1925, YIO, and b*) as well as colour difference were compared to scaling values (Z score) for each tooth sample as visual perception was used as a reference standard

for each experiment. This method directly evaluates the accuracy of the indices against perceptual evaluation of the set of stimuli.

2.9.3 STRESS Analysis & F-Test

Standard residual sum of squares (STRESS) is a statistical analysis method used evaluate if two indices or instruments are or are not statistically different with respect to a given set of visual data (García *et al.*, 2007; CIE 15, 2004). STRESS is used in multidimensional scaling techniques and has been used to quantify the performance of different colour difference equations (i.e. CIE94, CMC, and CIEDE2000) (García *et al.*, 2007; Kirchner and Dekker, 2011; Melgosa *et al.*, 2008). It is a restricted regression' model meaning data set should pass through the origin. STRESS analysis is unsuitable for data in which the smallest colour difference (i.e. index value) is not equal to zero (Kirchen and Dekker, 2011; CIE 217, 2016). An affine transform was applied to each data set so one sample for both the perceptual data and measurement data was equal to zero. An affine transform alters the data set by the following equation:

$$\boldsymbol{New\ Data = Ax + b} \qquad \textbf{Equation 2-9}$$

where A is the raw data set, x and b are variables of any value. With affine transforms the relations between the data point along the length of the line are preserved (Brannan *et al.,* 2011). STRESS is calculated using eqn. 2-10, where V is the visual scale values (Z scores), I is the computed index (i.e. WIO), and F1 is a scaling factor to minimize the STRESS.

$$\boldsymbol{STRESS = 100\left(\frac{\sum(I_i - F_1 V_i)^2}{\sum F_1^2 V_i^2}\right)^{\frac{1}{2}}} \qquad \textbf{Equation 2-10}$$

$$\boldsymbol{with\ F_1 = \frac{\sum I_i^2}{\sum I_i V_i}}$$

The major benefit to this statistical method is its simplicity and provides a lower and upper limit for a given metric. Results are in the range of 0-100 as they are presented as a percentage (García *et al.*, 2007). The greater the STRESS value represents a worse agreement between visual

perception and computed index. According to Kruskal (1964), a value of 0% has perfect correlation, excellent has a value of between 0.1 and 2.5, 5 is good, 10 is fair and anything above 20 has poor 'goodness of fit'. The aim is to achieve a STRESS value of zero. An F-test is a statistical method using the F-distribution to determine if two models a statistical significantly different or not based on a specific confidence interval. F-test for comparing two STRESS metrics is calculated with eqn. 2-11 (García *et al.*, 2007).

$$F = \frac{STRESS_A^2}{STRESS_B^2} \qquad \textit{Equation 2-11}$$

The confidence interval is developed on the critical value (Fc) with a confidence level used is of 95% with (N-1, N-1) degrees of freedom. This is the most commonly used confidence level and therefore used in this thesis (CIE 217, 2016; García *et al.,* 2007). The confidence interval is between the critical Fc value and 1/Fc, which can be found on a Fc critical table (Purdue, 2015). If the F value is within the specified confidence interval, the two indices are not statistically significant. If the F value is outside the specified confidence interval, the two indices are statistically significant. STRESS was calculated using MATLAB using the code in Appendix F.

3 Yellowness Index

3.1 Introduction

The global market for tooth whitening has grown significantly from $6 billion in 2020 and is projected to grow to $8 billion by 2026 (Stratview Research, 2020). New tooth-whitening approaches and products continue to be developed (i.e. trays, strips, Therasmile) as the mechanisms of 'whitening' have become more widely understood (Flucke, 2021). Many studies in dentistry investigate methods and indices for evaluating tooth 'whiteness' in order to assess the efficacy of these tooth whitening products.

The dental industry has associated the terms 'white' and 'yellow' as antonyms due to the physical processes of stain accumulation or aging (yellowing), while bleaching or brushing has been defined as 'whitening' (Pan and Westland, 2018; Paravina, 2008; Joiner and Luo, 2017). Due to this reason, whiteness has sometimes been expressed by the quantification of yellowness (Luo *et al.*, 2009). However, it is not proven if yellowness and whiteness are antonyms in practice (Pan and Westland, 2018).

Pan *et al.* (2018) conducted a cross-cultural study of perceptual whiteness and concluded that the concept of whiteness is consistent between different cultures, genders and age groups. This suggests that a single whiteness index could be used to measure changes in perceptual whiteness for these different groups. However, whiteness is not the only colour perceptual attribute of teeth that is of concern to dentists and patients. Teeth are yellowish in hue throughout the entire tooth colour gamut due to the underlying dentin (Hirschen, 2010; Joiner *et al.*, 2008a). During the bleaching or 'whitening' process it is actually the yellow chroma (i.e. stains) that is reduced. Therefore a yellowness index should be used for evaluating changes in tooth yellowness instead of 'whiteness'.

Multiple yellowness indices exist but were developed for use in different industries (i.e. paint, oils, plastics) meaning they may not be suitable for use in dentistry (ASTM E313-05, 2005; Hunter, 1981; Joiner and Luo, 2017; Mohan *et al.*, 2008). Despite this, some have been used in dental research (del Mar Pérez *et al.*, 2016; ASTM E313-05, 2005). There is no yellowness index that has been shown to be effective at correlating with changes in perceptual yellowness of teeth. The most commonly used method for evaluating changes in perceptual yellowness is the use of CIELAB parameter b* (Rubio *et al,* 2015). Studies have reported changes in lightness and yellowness, using changes in CIELAB parameters L* and b* to correlate with perceptual changes in lightness and yellowness respectively (Tao *et al.*, 2017b; Oliveira et al., 2015). With darker yellows (such as some tooth colours) b* is not recommended, but it is still used for evaluating changes in yellowness (Hunter, 1981). However, it is far from clear that b* can be used as a correlate to perceptual yellowness as neither of the cartesian coordinates a* and b* can indicate hue on their own. Therefore it is valuable to develop a yellowness index that is optimised for use in dentistry. In this experiment a yellowness index is developed for use in dentistry using data from existing unpublished data collected by Pan, Westland and Ellwood (2018) using Vita shade guide tabs.

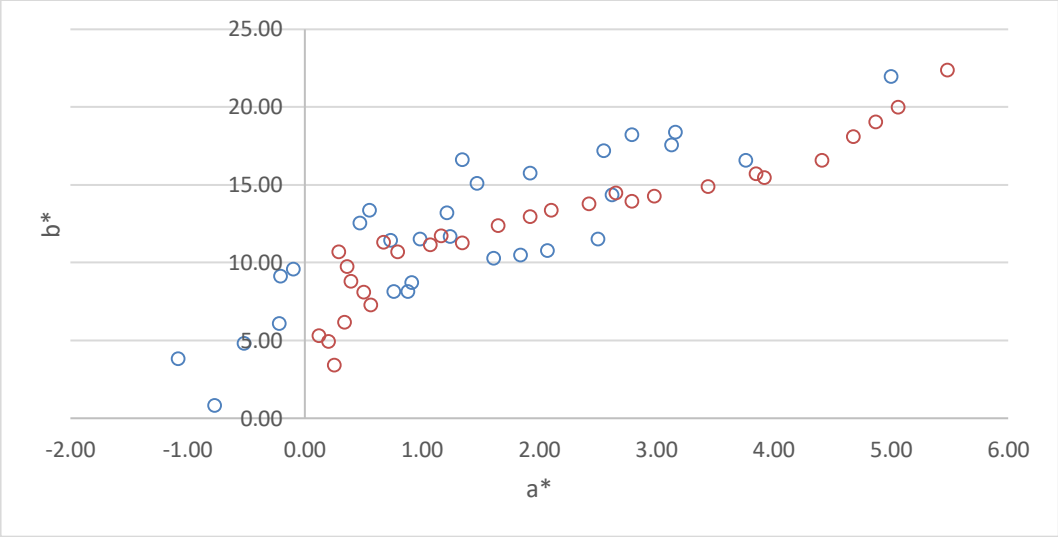

Figure 34: CIE a*–b* values of 29 Vita 3D Master Shade Guide (blue) and 29 Vita Extended Bleachedguide tabs (red)

The main limitation of shade guide tab samples is that there are correlations between the colorimetric values within the tooth colour gamut. In traditional 'whitening' methods such as bleaching or brushing as L* increases, b* decreases. In cases of ageing as L* decreases, b* increases; and as a* decreases, b* decreases. An example is shown in Figure 34. The correlation is due to the samples representing stages that are naturally described by bleaching and ageing within the tooth colour gamut.

With correlations in the data sets, multiple equations are able to fit the data with similar performance. While this is a problem, these equations may preform different when present with data that are not correlated. This is important in the evaluation of tooth whitening products as not all 'whitening' processes following the trend of as L* increased, b* decreases. Blue covarine toothpastes deposit a blue coating onto the surface of the teeth to make teeth visually appear 'whiter' for short-term period/up to a number of hours. With this coating as b*decreased, L* decreases (Collins *et al.*, 2008; Tao *et al.*, 2017b; Joiner *et al.*, 2008b). It is unclear if previously published indices or the new proposed yellowness index will perform in such cases. Five additional sets of un-correlated stimuli within the tooth colour gamut were generated. Due to the difficulty in creating

physical shade guides that were uncorrelated, the samples were digitally stimulated on a colour-calibrated screen for assessment. Un-correlated samples were essential in order to evaluate the robustness of these indices.

In the context of teeth, it is unclear what the relationship between perceptual whiteness and perceptual yellowness is. In dentistry, it might be expected that as perceptual whiteness increases, perceptual yellowness decreased due to the mechanisms of bleaching (i.e. reverse for aging). But to what extent are these concept antonyms, particularly within the constraints of tooth colour gamut? In order to address this question and to develop a yellowness index for use in dentistry, a psychophysical scaling experiment was conducted to measure visual yellowness for a set of samples. A yellowness index was developed based on these psychophysical data and was validated using data from a second experiment where participants viewed 5 sets of digital simulated teeth on a colour-calibrated display.

3.2 Psychophysical Experiment

In a study of perceptual whiteness conducted by Pan, Westland and Ellwood (2018), a psychophysical experiment was carried out where participants were asked to rank a set of shade guide tabs in order or whiteness. During the study they also asked participants to rank the same samples based on decreasing perceptual yellowness. Their unpublished yellowness data was analysed in this chapter. In their study, participants were asked to partake in an Ishihara Test to evaluate colour blindness. Colour-blind participants were excluded from the study. Participants were asked to rank the samples based on yellowness. The study was conducted in 5 countries: United Kingdom, India, Brazil, China, and United States. There were 100 participants from each country with 25 young males, 25 young females, 25 old females and 25 old males. The young participants were aged between 18 and 30 years old while the

participants categorised as old were between the ages of 30-60 years old. Each participant was asked to rank each of 58 shade guide tabs in order of yellowness in a lighting cabinet with D65 lighting arranging the samples from most yellow to least yellow.

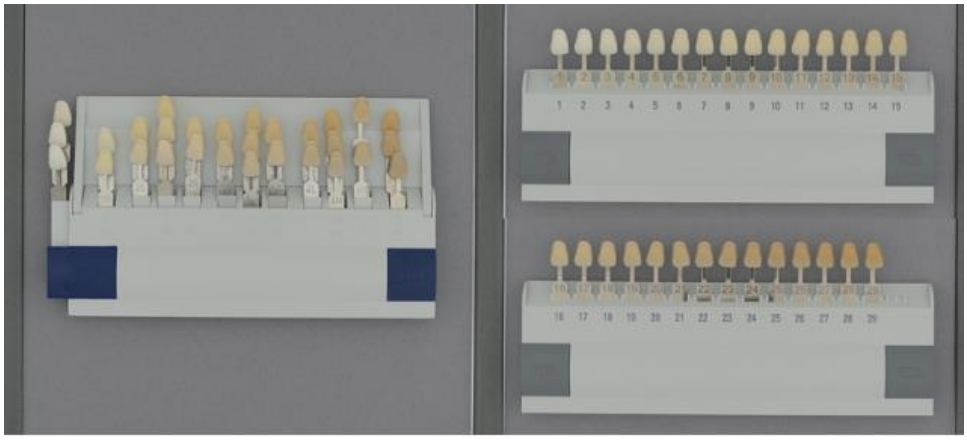

Figure 35: Vita Shade tabs. Vita Toothguide 3D master on the left and the custom made extended Vita Bleachedguide 3D master on the right

The set of 58 VITA shade tabs used in this study consisted of 29 VITA Toothguide 3D Master tabs and 29 custom-made shade tabs that extended the VITA Bleachedguide 3D Master (Figure 35). The Bleachedguide 3D master included 15 shade tabs. However, the custom-made extended Bleachedguide includes shades in between each normal shade tab in the guide to produce 29 total shades.

The samples were measured with a Konica Minolta CM-2600D. This data was collected and published by Pan, Westland and Ellwood (2018), but was available for use in this study.

3.3 New yellowness metric, YIO

The new yellowness index YI metric is based on the same generic form as WIC and WIO as shown in Eqn. 3-1 (Luo *et al.,* 2009).

$$\boldsymbol{YI = Y + p(x_n - x) + q(y_n - y)} \qquad \textbf{Equation 3-1}$$

The coefficients x_n and y_n are the chromaticity coordinates for D65 illumination (section 2.8). The values of Y, x and y values refer to measurements from the 58 tabs. The coefficients p and q were optimised to maximise the r^2 value (closest to r^2=1) between YI and the perceptual yellowness values (Z-scores) for the 58 samples that were ranked by 500 participants. The optimisation was performed using the Solver algorithm in Microsoft excel. The new optimised yellowness metric, YIO, resulted in Eqn. 3-2.

$$\boldsymbol{YIO = -Y - 851.716(x_n - x) - 436.962(y_n - y)} \qquad \textbf{\textit{Equation 3-2}}$$

3.4 Validation Experiment

This experiment was approved by the University of Leeds Ethics committee (LTDESN-084) (Appendix G). The validation experiment was collected based on five related psychophysical experiments in which participants evaluated digital tooth images based on perceptual yellowness.

3.4.1 Temporal Stability of Display

An iiyama display was used for the experiment. The display had a set contrast and brightness of 75 cdm^{-2}, and the colour temperature of the screen set at 6500K. To measure the temporal stability of the display, a Konica Minolta Spectroradiometer set up following section 2.2.3, measured a white screen every 30 seconds for 1.5 hours.

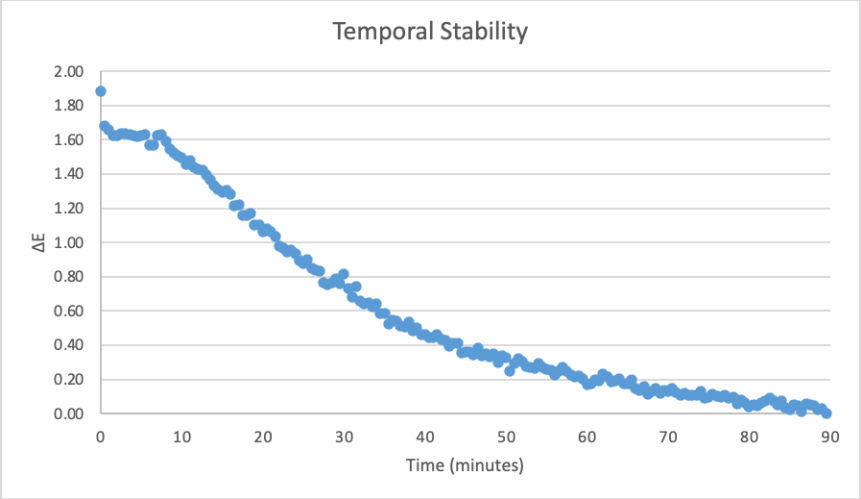

Figure 36: Temporal Stability of iiyama Display

Figure 36 displays the temporal stability of the iiyama display measured by the spectroradiometer for 90 minutes. The colour difference, ΔEab, was calculated by computing the difference between each measurement and the very last measurement (taken after 1.5 hours). Theoretically, the device reaches temporal stability when the ΔEab reaches 0. However, this will only happen with the last measurement as the device inherently has measurement variability. The spectroradiometer has inter-instrument variability of ±2% for luminance measurements, meaning the ΔE will never be 0 (Konica Minolta, 2007). In

Figure 36 displays the temporal stability of the iiyama display. Prior to the main experiment taking place, the display was turned on and left for 60 minutes to reach temporal stability, which has a Δ E=0.17.

3.4.2 Samples

Digital images were chosen for this study versus shade guide tabs in order to ensure uncorrelated samples within the tooth colour gamut. Digitally simulated teeth have been used in previous studies (Höfel *et al.*, 2007). The L* and b* values of the samples are shown in Figure 5. If we take all of these samples together, we will find that L* and b* are strongly correlated (this is inherent in the gamut of tooth colour). However, by separating the data into 5 sets we can see that, within each set, there is

little to no correlation. This provides a robust test of the yellowness index. Pan, Westland, and Ellwood (2018) provided the base images from which the images in Figure 37 (right) were constructed.

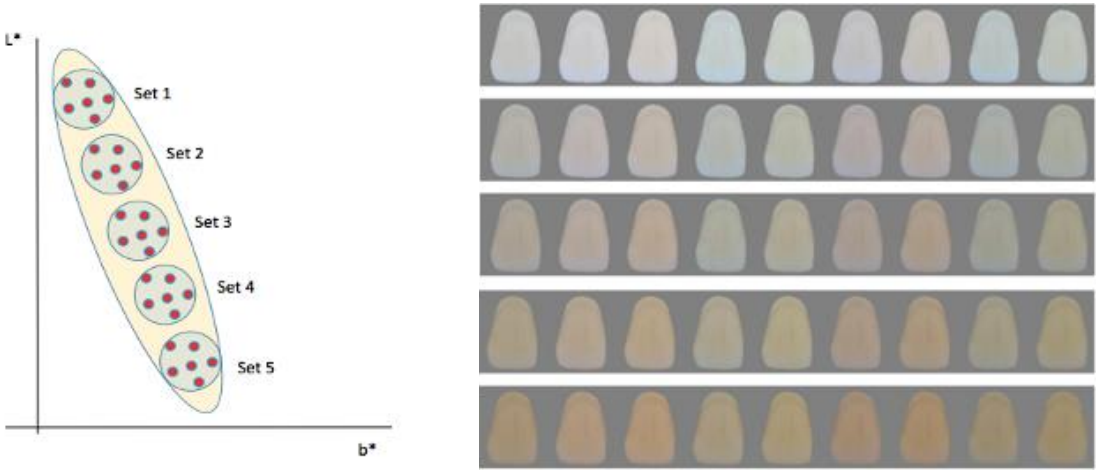

Figure 37: Image of method to de-correlate the tooth image (left) and the resulting 5 sets of 9 stimuli developed for evaluation (right)[1].

The 45 images were broken up into 5 sets of 9 tooth images which span the tooth colour gamut. For each tooth image, the RGB values of the baseline tooth image was digitally modified to provide specific CIELAB values for each image which were completely uncorrelated from each other in colour space. Each set of 9 stimuli was based on a different colour centre (Table 5).

Table 5: Pan et al. (2018) target CIELAB values for the five colour centres

Colour Centre	L*	a*	b*
Set 1	78.9	-0.8	0.8
Set 2	70.9	0.2	4.0
Set 3	66.4	0.9	8.1
Set 4	64.8	1.2	13.2
Set 5	61.0	2.9	17.4

[1] The images may appear differently in this document than when the images were displayed on a calibrated iiyama display

Image Number in each set	L*	a*	b*
1	Colour centre		
2	+2ΔL	+2Δa	-2Δb
3	+2ΔL	+2Δa	+2Δb
4	+2ΔL	-2Δa	-2Δb
5	+2ΔL	-2Δa	+2Δb
6	-2ΔL	+2Δa	-2Δb
7	-2ΔL	+2Δa	+2Δb
8	-2ΔL	-2Δa	-2Δb
9	-2ΔL	-2Δa	+2Δb

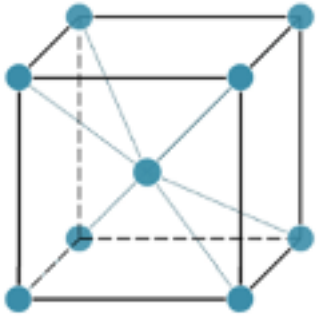

Figure 38: Distribution of 9 stimuli for one colour centre in CIELAB space and the changes of CIELAB values for each image of each set from the colour centre in a given set

The additional 8 other samples in each set were approximately at the vertices of a cube arranged around the colour centre, each being 2 CIELAB units away from the centre in each of the L*, a*, b* directions shown in Figure 38. When displayed on the colour calibrated iiyama display the values of the additional 8 samples are not exactly +/- CIELAB values like the table in Figure 38.

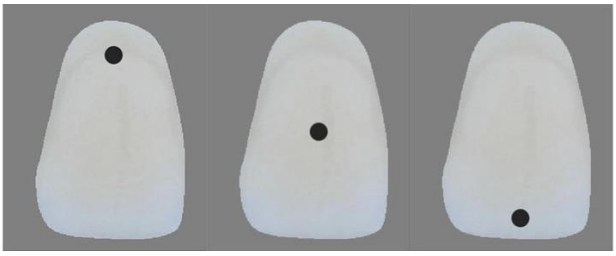

Figure 39: Tooth image with aperture of spectroradiometer in which measurements were collected at the gingival (left), central (middle), and incisal (right) location. This is what the image would look like through the spectroradiometer aperture lens

Each tooth image was measured at the gingival, central, and incisal location with a spectroradiometer following methodology section 2.2, as shown in Figure 39. The colour centres when displayed on the iiyama screen are shown in Table 6 (the measured values in Table 6 can be compared to the target values in Table 5). The CIELAB values for all the samples when displayed on the iiyama screen are shown in Appendix H.

Table 6: Values of colour centres on colour calibrated iiyama display screen

Sample Set	L*	a*	b*
Set 1	81.01	-2.08	0.25
Set 2	74.32	-2.09	3.15
Set 3	70.31	-2.03	6.96
Set 4	68.78	-2.14	11.53
Set 5	65.33	-1.23	15.42

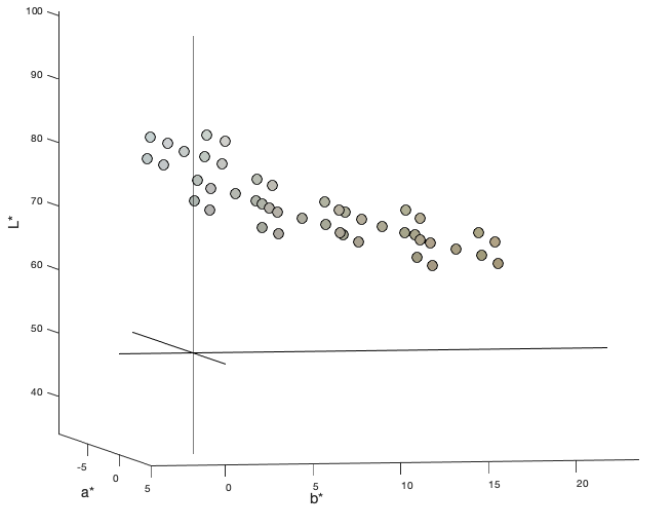

Figure 40: Distribution of 45 samples in CIELAB space and visual representation of colour sample when displayed on iiyama screen

The resulting distribution of the 45 samples is shown in Figure 40, the rough ‘cube’ structure for each data set can be seen. The visual representation of the samples from CIELAB measurements are shown in Figure 40. For each of the 5 experiments, the 9 digitally simulated teeth were displayed using a graphical-user interface (GUI) written in MATLAB software on a colour calibrated iiyama display. One set of 9 samples are presented at a time on the GUI a grey background which had an average CIELAB value of L*=57.12, a*=-1.36, b*=1.00.

3.4.3 Psychophysical Experiment

A total of 40 volunteers participated in each of the experiments. These participants were no screened for colour-blindness in order to obtain results from 'regular' observers. An email was sent asking for participants to volunteer their time. Each participant was given a small financial reward to compensate them. Those available arranged a time via an online scheduling platform. Prior to participation, volunteers were given a participant information sheet and asked to sign a consent form (Appendix I-Appendix J). There were 27 females and 13 males between 22 and 55 years of age. All participants had to be over the age of 18 in order to given consent to participate. Participants were asked to sit in front of a calibrated iiyama computer screen. One set of teeth would appear at a time to be arranged by the participant. To ensure complete randomization of the data, each set of teeth were presented in a different order each time for example set one was not always the first set seen by the participant.

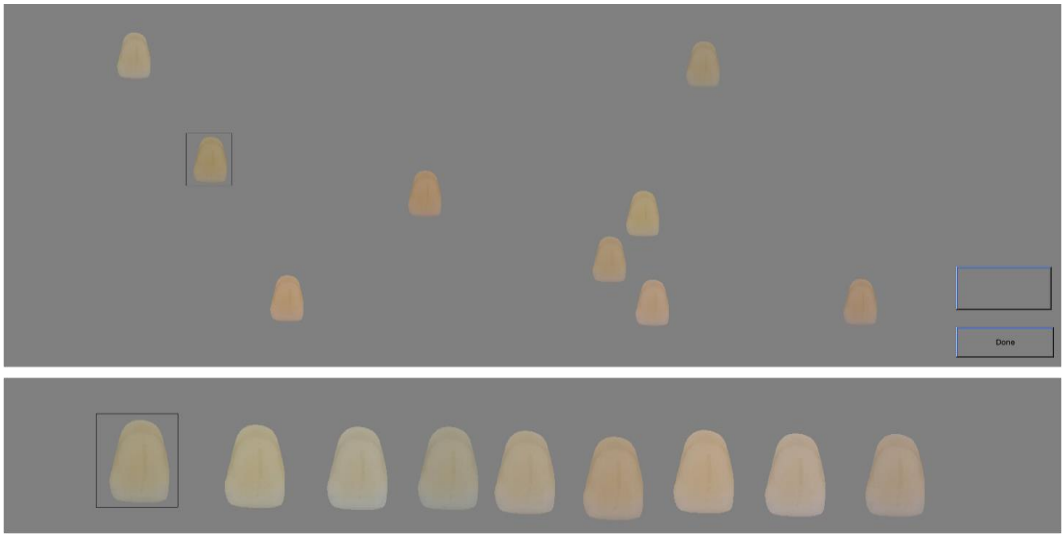

Figure 41: Before (top) arrangement and after (bottom) ranking of a set of 9 digital tooth images based on perceptually yellowness

The tooth samples appeared in random locations and participants were asked to use the mouse to arrange the samples, ranking them based on decreasing yellowness from left to right (Figure 41). Participants were asked to ignore the black box around one of the tooth samples as that was

a programming error in the GUI that could not be fixed. They repeated this procedure 5 times until each set is completed. The rank order data was recorded for each participant and for each set. Interval scales were calculated for each set separately using methodology section 2.4.1. Interval scale values are shown in Table 7. Note that the Z scores in Table 3 are relative within each set. However, Z scores cannot be compared between sets because no sample on one set was ever visually compared with any sample in any other set.

Table 7: Interval Scales (Z-scores) for each set of images based on 40 perceptual rankings participants

Image Number in each Set	Set 1	Set 2	Set 3	Set 4	Set 5
1	-0.06	0.08	-0.19	-0.04	-0.08
2	-0.45	-0.68	-0.44	-0.48	-0.53
3	0.45	0.19	0.34	0.20	0.22
4	-0.87	-0.75	-0.58	-0.72	-0.64
5	-0.05	-0.10	0.06	0.06	0.02
6	0.14	0.22	-0.14	0.10	0.13
7	0.72	0.81	0.83	0.65	0.65
8	-0.35	-0.13	-0.34	-0.29	-0.21
9	0.45	0.34	0.49	0.50	0.44

3.5 Assessment of Yellowness Indices

Three statistical methods are used to quantify the agreement between the candidate yellowness indices and the psychophysically derived yellowness interval scale values (Z-scores). Coefficient of determination (r^2) and percent wrong decision (%WD) were used as they have been used in previous dental studies and allow for easy comparison (Luo *et al.,* 2009; del Mar Perez *et al.,* 2016). STRESS analysis is introduced is this thesis for use in dentistry as an alternative method to evaluate statical differences in performance of the different indices that are used (Kirchner and Dekker, 2011; Sullivan *et al.*, 2019). The three candidate yellowness indices evaluated in this study were WIO (Eqn. 1-5), b* and a new yellowness equation YIO (Eq. 3.2). The metric b* is used to evaluate yellowness in

this study as previous studies have used b* to measure tooth colour changes during tooth-whitening procedures (Tao et al., 2017b; Oliveira et al., 2015; Kim et al, 2000). The use of STRESS implies a restricted regression. Therefore the data for the 58 samples were subject to an affine transform so that sample 0M1 was equal to zero for the perceptual data and indices prior to STRESS analysis. For the 45 digital simulated samples, an affine transform was applied to the data so that each of the colour centres for the perceptual data and indices was zero prior to STRESS analysis. Note that an affine transform has no effect on calculations of r^2 or %WD. Statistical data between index values and perceptual yellowness (Z-scores) were calculated using methods in section 2.9. WIO is a whiteness index in which a higher value represents increased whiteness. For this study, lower values of WIO will be used to represent increased 'yellowness'.

3.6 Results

The r^2, %WD, and STRESS values for the three candidate yellowness indices on the training set of 58 shade guide samples are shown in Table 8. Higher values of %WD and STRESS indicate worse agreement. The closer to 1 the r^2 value, the better the agreement.

Table 8: Performance of the indices on the 58-tab data that was derived from measurements collected by Pan et al., 2018

Index	r^2	%WD	STRESS
WIO	0.97	4.84	7.45
YIO	0.97	4.96	7.01
b*	0.91	9.68	11.57

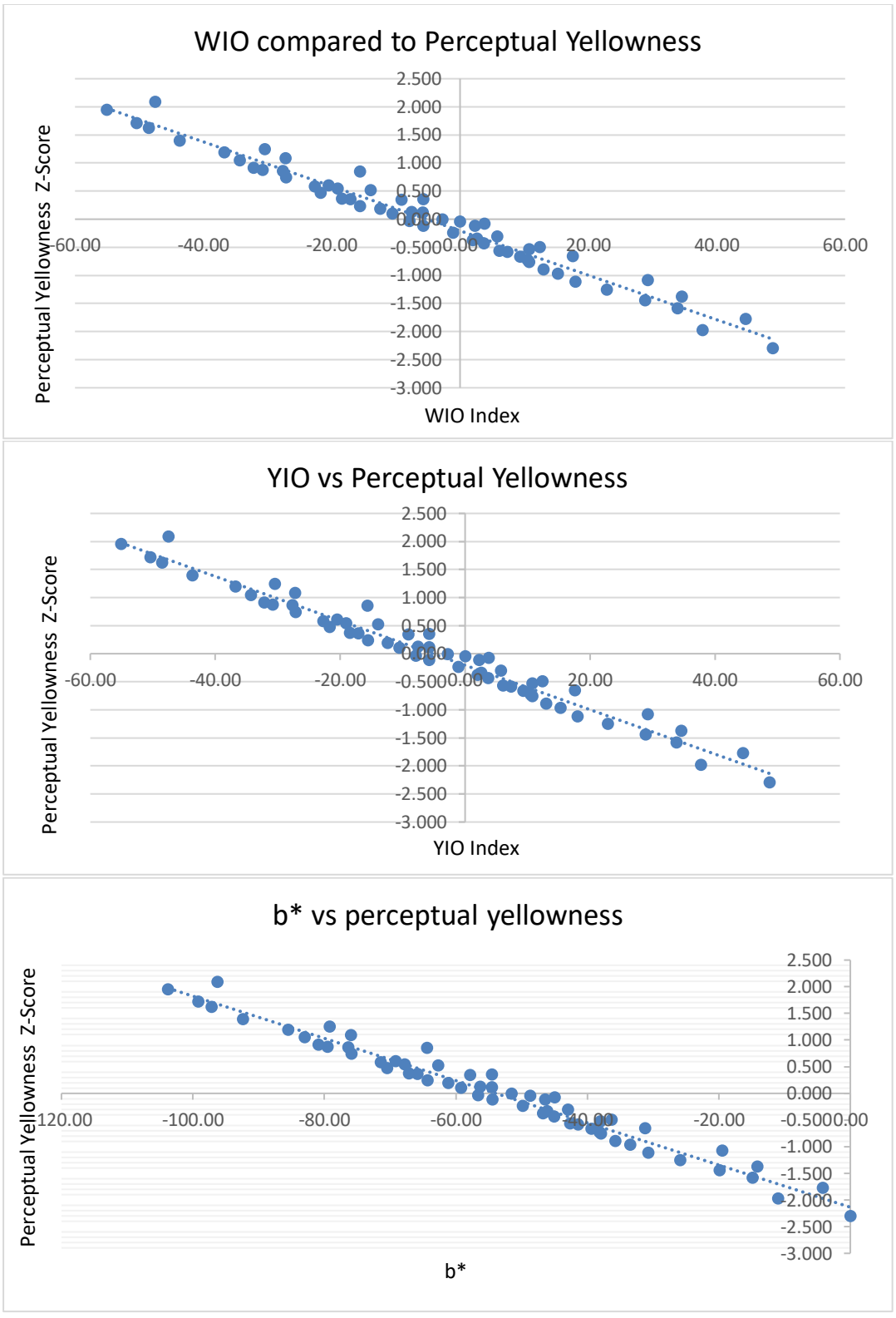

Figure 42: Correlation between the three indices and the perceptual yellowness data for the 58 samples. Data are shown for WIO (upper), YIO (middle) and b* (lower) with r^2 values of 0.97, 0.97 and 0.91 respectively.

Figure 42 displays a graphical representation of the variance between perceptual yellowness and candidate index for the 58 samples. It is clear from both visual representation (and the r^2 values) that b* is less suitable than either WIO or YIO for evaluating changes in perceptual yellowness. This is further confirmed with higher %WD (9.68) and STRESS scores (11.57) than WIO and YIO (Table 8). A STRESS score of 11.57 is considered to have fair correlation between the data. Both WIO and YIO have identical r^2, and very similar %WD and STRESS values. According to STRESS, YIO performs slightly better than WIO.

Table 9: Squared STRESS ratios comparing candidate yellowness indices

Index Comparison	Squared STRESS ratio
YIO vs. b*	0.37
WIO vs. b*	0.41
WIO vs. YIO	1.13

The squared STRESS ratios are shown in Table 9. The critical F value for F_c (0.975,57,57) is 1.73 giving confidence interval of 0.58-1.73 (Purdue, 2015). If the squared STRESS ratios fall outside of this confidence interval it indicates statistical difference. Based on the results shown in Table 9, b* is statistically different from YIO and WIO. However, WIO and YIO formula are statistically indistinguishable. Although YIO was optimised based on these data, it is interesting that WIO performs just as well in evaluating perceptual yellowness.

Table 10: Performance of r^2 of the indices on the five sets of digitally simulated data from 40 participants

	Set 1	Set 2	Set 3	Set 4	Set 5	**Mean**
WIO	0.81	0.95	0.81	0.92	0.94	**0.88**
YIO	0.79	0.94	0.84	0.94	0.96	**0.90**
b*	0.55	0.38	0.70	0.57	0.58	**0.56**

Table 11: %WD results of the indices on the five sets of digitally simulated data from 40 participants

	Set 1	Set 2	Set 3	Set 4	Set 5	**Mean**
WIO	11.11	2.78	11.11	8.33	8.33	**8.33**
YIO	16.67	2.78	11.11	8.33	2.78	**8.33**
b*	38.89	38.89	22.22	33.33	22.22	**31.11**

Table 12: STRESS results of the indices on the five sets of digitally simulated data from 40 participants

	Set 1	Set 2	Set 3	Set 4	Set 5	**Mean**
WIO	47.41	28.93	58.10	30.98	32.56	**39.59**
YIO	47.93	28.34	56.48	26.28	28.83	**37.57**
b*	66.79	78.67	65.67	65.40	65.89	**68.48**

Table 10, Table 11, and Table 12 display the r^2, %WD and STRESS results as well as the mean scores for the five sets of uncorrelated digitally simulated images. When evaluating the indices against the five sets of uncorrelated digitally-simulated images, b* producing the lowest r^2 value with a mean of 0.56, shown in Table 10. While WIO displayed high correlation with perceptual yellowness with an r^2 mean value of 0.88, YIO produced a stronger correlation between index and perceptual yellowness with a mean r^2 score of 0.90. This is to be expected since YIO was optimised for evaluating yellowness. It is interesting to note that WIO and YIO indices had better correlation with perceptual results for sets 2, 4 and 5 than sets 1 and 3. Table 11 show the %WD results in which b* produced higher mean value of 31.11, compared to WIO and YIO which had the same mean value of 8.33. Similar to r^2, WIO and YIO %WD data correlated better with sets 2, 4 and 5 produce better results than 1 and 3. It is unsurprising that YIO produces a higher %WD of 16.67 for set one, which has least yellow CIELAB values. According to STRESS values shown in in Table 12, all three sets of the 5 sets have mean values of 37.37 or above. YIO produced the lowest STRESS values overall indicating better correlation with perceptual yellowness than WIO and b*.

Table 13: Squared STRESS ratios comparing candidate yellowness indices for the 5 sets of digitally simulated images

Index Comparison	Set 1	Set 2	Set 3	Set 4	Set 5	**Mean**
YIO vs. b*	0.52	0.13	0.74	0.16	0.19	**0.30**
WIO vs. b*	0.50	0.14	0.78	0.22	0.24	**0.33**
WIO vs. YIO	0.98	1.04	1.06	1.39	1.28	**1.11**

The squared STRESS ratios calculated are shown in Table 13. The critical F value for F_c (0.975,8,8) is 4.43 giving a confidence interval of 0.23-4.43. Comparing the mean values, none of the indices are statistically different. This large interval range could be due to the small sample size. Evaluating the squared STRESS by set change the YIO and b* are statistically different for set 2, 4 and 5. WIO is only statistically different from b* for set 2 and 4. For all 5 sets, WIO and YIO indices are statistically indistinguishable.

All three statistical methods confirm that b* is an inadequate method for analysis of evaluating perceptual yellowness. Based on STRESS alone, YIO performs better than WIO and b*. From the r^2 and %WD, both WIO and YIO are suitable for evaluating perceptual yellowness of samples that are uncorrelated. However, YIO did perform slightly better than WIO according to r^2 and %WD for evaluating perceptual yellowness. Based on STRESS, YIO performed slightly better than WIO, but are statistically indistinguishable.

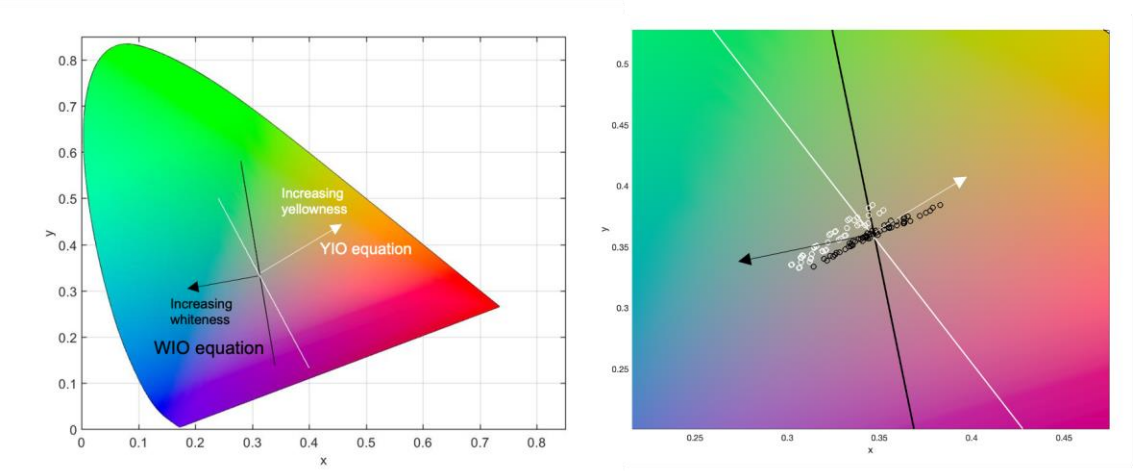

Figure 43: Iso-whiteness (black line) and Iso-yellowness (white line) for the WIO and YIO equations respectively on a chromaticity diagram (left). On the right is the diagram with iso-whiteness and iso-yellowness with the 58 shade guide tab data (black circles)

Figure 43 graphically illustrates the difference between the WIO and YIO equations using the concept of iso-whiteness and iso-yellowness lines in a CIE chromaticity diagram. In this diagram all of the colours along the white line have the same perceptual yellowness as predicted by Eqn. 3-1. All the colours along the black line have the same perceptual whiteness as predicted by Eqn. 1-5 (WIO). A set of imaginary lines parallel to the white line would also each have the same iso-yellowness. Yellowness increases perpendicular to these lines towards the yellow region of colour space as denoted by the white line with arrow. The black line with arrow shows that increasing in whiteness according to WIO is from an orange region to cyan region of the chromaticity diagram. Whereas increasing yellowness according to YIO is from the blue to yellow regions of the diagram. This would suggest that although tooth whiteness and tooth yellowness are highly related concepts, yellowness is not simply the antonym of whiteness. On the right of Figure 42, the 58 samples measurement by a spectrophotometer (black circles) and the 45 digitally simulated images measured by a spectroradiometer (white circles) are placed on the chromaticity diagram with the iso-whiteness and iso-yellowness lines. Neither the 58 samples or the 45 uncorrelated samples

fall perfectly align with iso-yellowness or iso-whiteness. Based on STRESS and %WD, YIO and WIO perform equally as well. It is clear from this graph that other hues slightly influence tooth colour other than just yellowness and whiteness.

3.7 Discussion

In terms of r^2, all three indices showed good correlation with perceptual yellowness for the 58 shade guide tabs. WIO and YIO produced identical r^2 values that were a little higher than b*. Both indices also produced similar %WD values that were lower than b*, indicating better performance. Based on STRESS analysis, b* is an unsuitable metric for predicting changes in perceptual yellowness. The results suggest perceptual yellowness of teeth is not simply a correlate of CIELAB b*. It is interesting that the STRESS results found that WIO and YIO were statistically indistinguishable. The strong performance of WIO relative to YIO is unexpected as YIO was developed on this test set. The performance of WIO could be due to the fact that index was developed using data from VITA shade guide tabs as well (Luo *et al.,* 2009).

The performance of the three indices on the 5 sets of digitally simulated tooth samples is weaker than on the 58 shade guide tabs. Similar to the results of the 58 samples, WIO and YIO performed better than b*. This suggests that b* should not be used for evaluating changes in perceptual yellowness. The digitally simulated samples constitute a much more rigorous evaluation for the equations. For these samples, as b* decreases, L* may decrease, increase or remain constant. However, for sets of shade guide tabs, the colorimetric values are highly correlated. Changes in colour between tabs move along the gamut of natural tooth colour changes. For example, when individuals age their tooth colour L* will decrease and b* will correspondingly increase (Xiao *et al.*, 2007). Alternatively, dental bleaching procedures increase L* and decrease b* (Pan and Westland, 2018). The correlation of the physical samples reduces the degrees of

freedom in the data, the number of components in data that can vary while still yielding a given population a value for characteristics. If tooth colour has three values (L*, b*, and a*) the degree of freedom would total to two. With L* and b* being heavily correlated, the degree of freedom in the data is reduced to one. The larger the degrees of freedom, the more likely the results will be statistically significant (Howitt and Cramer, 2014). With one degree of freedom, this enables the ability to have multiple equations fit the data equally. Using the same set of 500 data and an updated version of excel solver, a different index can be produced:

$$\boldsymbol{Yellowness} = \ -\boldsymbol{Y} - \mathbf{801.845}(\boldsymbol{x_n} - \boldsymbol{x}) - \mathbf{630.383}(\boldsymbol{y_n} - \boldsymbol{y}) \quad \textbf{\textit{Equation 3-3}}$$

However, the robustness of each index developed will vary. Robustness of an index is important to improve accuracy of tooth colour measurement as well as evaluate blue covarine toothpastes as b* decreases, L* decreases (Collins *et al.*, 2008; Tao *et al.,* 2017b). The candidate indices were evaluated against uncorrelated data for this reason. Based on STRESS analysis, none of the candidate metrics are suitable for evaluating changes in yellowness. STRESS found that there was no statistical difference between YIO and b* for two of the 5 sets and no statistical difference between WIO and b* for three of the 5 sets of validation data. It is interesting that the squared STRESS shows that there no statistical difference between the three metrics for a couple of the sets of validation data considering b* produced poor r^2 and %WD results for all 5 sets. The STRESS results could be affected by the small sample size (n=9) for the 5 sets of data. However, a small set was required to ensure the samples were uncorrelated from each other within the tooth gamut colour space.

The yellowness index was optimised on the 58 samples. It would be expected that the r^2 value should be 1 and the STRESS should be close to 0. However, the r^2 value was 0.97 and STRESS value of 7.01, which is

considered between good and fair correlation (Kruskal, 1964). This could be due to the fact that the index used data from measurements collected from a diffuse/0° spectrophotometer in order to aid in optimising the equation to correlate to visual perception. Spectrophotometers are known to not have the most 'accurate' measurements due to edge-loss (Bolt *et al.,* 1994). However, WIO was optimised using measurement data from a spectroradiometer with 45°/0° illumination and the index performed just as well on the test data (Luo *et al.,* 2009). With multiple devices available, using data from a digital camera or Easyshade could produce a more robust index. However, it is unclear what device is best as there is no gold standard instrument to use in dentistry.

A different method that might develop a more robust index is to use principal component analysis instead of Excel Solver. It is a statistical method which identifies major relationships in complex data (Hess and Hess, 2018; Abdi and Williams, 2010). It is a useful metric for understanding correlated relationships by creating artificial uncorrelated variables optimized to maximise the variation in a set of data for each component (Hess and Hess, 2018; Abdi and Williams, 2010; Westland *et al.,* 2012). This method might be able to isolate the different attributes of tooth colour that are not accounted for when using excel solver.

White is an attribute of colour perception by which an object's colour is judged to approach the 'perfect' white (i.e. achromatic) (ASTM E313-05, 2005; Pérez et al., 2016; Hirschen, 2010). The tooth colour gamut is not devoid of hue.

Figure 44: Visual display of 5 of the 52 tooth coloured samples

Hirschen (2010) states that lightness, chroma and hue are evaluated prior to observer assessing which of these sensations is to be called 'white', especially with samples like in Figure 44 without a reference 'white'. Yellowness is an easier attribute to evaluate for teeth as even the 'whitest' sample has hue along the blue-yellow axis with b* values ranging from -6-48.6 (Cho *et al.*, 2007; Yamanel *et al.*, 2010; Jarad *et al.*, 2005; Mahn *et al.*, 2021; O'Brien *et al.*, 1990; Yuan *et al.*, 2007; Paravina *et al.*, 2006). The results indicate that perceptual yellowness produces better correlation to instrumental measurements than perceptual whiteness for all three metrics (WIO, YIO and WID). It could be argued that perceptual whiteness should not even be evaluated.

It is interesting that indices YIO and WIO preform equally as well on the 58 data set and the 45 uncorrelated data. Based on Figure 43, it can be seen that whiteness and yellowness are not perfect antonyms. Neither of these attributes displayed by iso-whiteness and iso-yellowness perfectly reflect tooth colour as it can be seen that red hues have an influence on tooth colour. These red hues might not be taken into account in the development of these indices and hence their poor correlation. If the perceptual data for the optimisation of an index was not based on a 'specific' hue but based on assessment of health or attractiveness the index, other colour characteristics will be included into the index improving correlation.

3.8 Conclusion

This study has provided two new psychophysical experiments to assess perceptual yellowness of teeth. A new yellowness index, YIO was

optimised for tooth colour. WIO and b* were also evaluated as predictors of perceptual tooth yellowness. In both the test set and validationn set of data, results found that b* is not an appropriate metric for assessing yellowness. YIO performed slightly better for the validation data sets than WIO. However, WIO and YIO according to r^2 and %WD are suitable for measuring changes in perceptual yellowness and this was further confirmed by STRESS which found the indices statistically indistinguishable.

This yellowness index has been published in the Journal of Dentistry (Sullivan *et al.,* 2019). Colgate-Palmolive has used YIO index in clinical trials evaluating the efficacy of a whitening toothpaste and created an advertising claim “reverses up to 15 years of discoloration” for the product Optic White Renewal Toothpaste based on results from changes in YIO (Hogan, 2021; Colgate-Palmolive Company, 2021).

4 Edge-Loss

4.1 Introduction

Edge-loss is the coined term for light ‘lost’ in colour measurement of translucent materials (i.e. skin, teeth, apple flesh, beeswax, cocoa butter, marble). This phenomenon presents a challenge for ‘true’ colour measurement of translucent materials as light scatters through the dental material exiting in a location outside the measurement area potentially producing measurements that are lower (darker) in comparison to how the tooth may perceptually appear. While lost light is inevitable for translucent materials, there is no consistent definition for edge-loss. Bolt and colleagues (1994) defined edge-loss as un-measured reflected light; they specifically stated that light exiting the edges or through the back of the material are not counted as edge-loss, even though that light is not measured by the detector. Other researchers defined the light that exits the edges and back as edge-loss (Ahn and Lee, 2008; Lee *et al.*, 2011). Edge-loss has also been described as displaced light from lateral scattering, specifically excluding light which is transmitted through the back of the tooth (Johnston *et al.*, 1996; Pop-Ciutrila *et al.*, 2016). However, the second two definitions describe transmittance; light that enters one boundary passing through a medium and exiting a different boundary (Knight, 2008; Al-Azzawi, 2007). Edge-loss in this thesis will defined as unmeasured light which is ‘lost’ from the standard illuminant due to transmittance and unmeasured reflectance (Gevaux *et al.*, 2020). The amount of edge-loss is influenced by device, illumination area, measurement area, as well as the degree of translucency and size of sample (Lee *et al.*, 2004; Kim *et al.*, 2008). This experiment explores the differences in the amount of edge-loss that occurs when using different instruments when samples vary in size, finish, and translucency.

Many researchers have stated that contact-based devices are subject to edge-loss, whereas non-contact based devices are void of edge-loss

(Joiner and Luo, 2017; Joiner, 2004; Pop-Ciutrila *et al.,* 2016). These claims stem from a study conducted in 1994 by Bolt and colleagues where 27 extracted incisors were measured using a 45°/0° spectrophotometer with varying measurement window size (i.e. target mask) of 3mm, 4mm, and 5mm, in which the illumination and measurement area were the same (i.e. illumination area 3mm, measurement area 3mm). The samples were also measured with a spectroradiometer. Even with a measurement area of 0.5mm x 2mm, the study claims the window size as infinite since the device does not make contact with the sample allowing illumination of entire sample. Results indicated as L* values decreased, a* values shifted more towards green (-a values), and b* values shifting more towards blue (-b values) as the window size decreased. While their results are valid, the conclusions drawn are based on their narrow definition of edge-loss; un-measured reflected light. This does not fully describe what is happening on a physical level.

Teeth are a 3D object comprised of translucent layers (i.e. turbid material) of enamel and dentine which reflect, absorb, scatter, and transmit light (Da Silva *et al.*, 2008; Wang *et al.*, 2013). When describing infinite sized samples, light only exits via reflectance or transmittance through the back on the sample. However, teeth have a finite size meaning light scatters though the material exiting at different points than it enters (Gevaux *et al.*, 2020). Light is able to enter and exit through the front, back and sides. According to Hemholtz reciprocity principle, light travelling through a non-photoluminescent sample from any point A to any point B will follow the same pathway from B to A with equal intensity when stationary (Clarke and Parry, 1985; Mishchenko *et al.*, 2002).

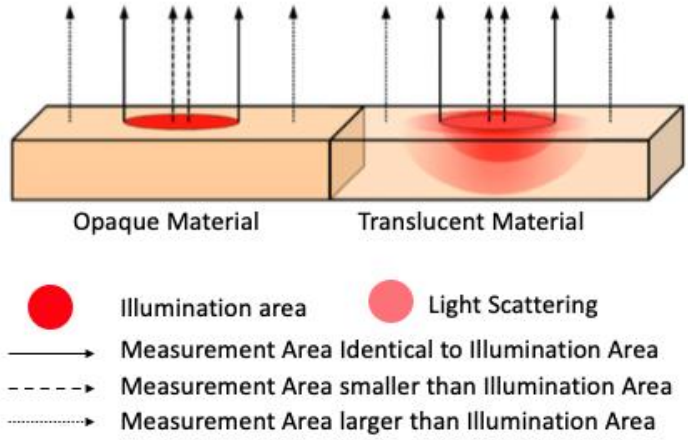

Figure 45 Gevaux et al. (2020) diagram of their theory about edge-loss

Figure 45 shows an image from Gevaux *et al.* (2020) study in which two different materials are illuminated in a finite area with varying measurement areas. With opaque materials/high scattering, little sub-surface scattering occurs meaning there is no edge-loss. When the measurement area is larger than the illuminated area when measuring a translucent material, the sub-surface scattering flux that emerges outside the illumination area is also measured. If the measurement area is smaller than the illumination area, they argue that the loss is perfectly compensated by light that has entered into the surrounding area which travels into the measurement area to be detected. This is a detailed explanation of the standard methods that exist for translucent materials which is to ensure as much possible light illuminates the sample in order to avoid a large amount of light exiting the sample as transmittance (i.e. back or sides). To do this, either illuminate a very small central area of the sample and view it with a large specimen port in order to capture as much light as possible. This method is difficult due to the already small size of teeth. The alternative is to uniformly illuminate a very large area of the specimen (i.e. entire sample) and measure a small central portion of the sample (ASTM E1164-12, 2017). While this is easily done for non-contact based device, it is more difficult for contact based devices. Commercially,

most contact-based measurement devices (i.e. spectrophotometers) are designed with the illumination area larger than measurement area. The relationship between amount of light lost from the incident illuminant and gained from ambient light source differs between contact based devices and non-contact based devices effects the amount of overall edge-loss.

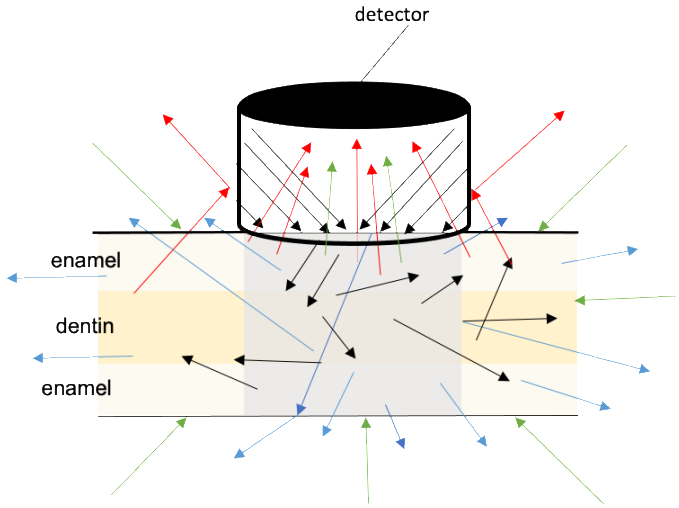

Figure 46: A schematic diagram of a contact-based device with a fixed aperture (i.e. window size) that has the same illumination area and measurement area

Contact-based devices intensely illuminate a sample in a finite area and detects light in a finite area. Figure 46 shows a schematic diagram of the different light interactions. The red arrows represent reflected light, blue arrows represent transmittance, green arrows are ambient light entering the sides and back of the sample (i.e. some of which is detected and compensates for the light lost from the standard illuminant), and the black arrows represent light from illumination and light scattered through the sample and the grey area displays how much of the sample is illuminated by standardised illuminant. The amount of light lost from the standard illuminant exiting the sample outside measurement area is greater than the amount of light travelling from the sides and back exiting within the measurement area producing lower L* values (darker). This is due to the target mask (i.e. window size) impedes the illuminating the entire sample with the same intensity and restricts both transmitted light and reflected light from being measured (i.e. transmittance and reflected light).

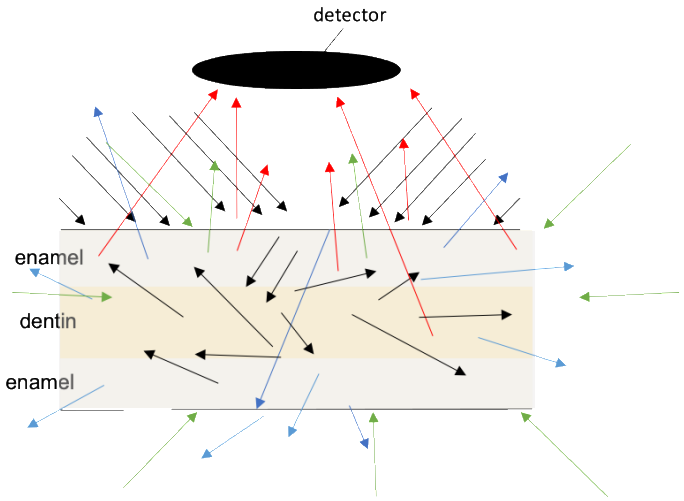

Figure 47: A schematic diagram of a non-contact based device

For non-contact based devices, the illumination source is separate from non-contact based devices allowing for teeth to be entirely illuminated. In addition, these devices have no target masks. Meaning there is no mask to prevent light from reaching the detector resulting in higher L* values (lighter) (Figure 47). More specifically, a larger amount of light (i.e. transmitted or reflected) from the standard illuminant from outside the measurement area is able to compensate for the amount of light 'lost' from the standard illuminant through the edges and back of the tooth reducing the amount of overall edge-loss. This is due to the relationship between illumination area and measurement area, in which the illumination area is significantly larger than the measurement area.

Lee *et al.* (2004) conducted a study investigating the amount of 'edge-loss' that occurs in dental resins based on the ratio of illumination area to measurement area (i.e. viewing area). Five unfinished and five finished resin composites (12mm diameter x 2mm thickness) of A2 shade were measured with a spectrophotometer with illumination area/measurement areas of 3mm/3mm and 11mm/8mm. The L* values increased, a* values became more red and b* values became more yellow when switching target masks from 3mm/3mm to 11mm/8mm for both the finished and unfinished samples. There was a larger L* increase for unfinished samples

when increasing the size of the target mask than there was for finished samples. These results further confirm Bolt *et al.* (1994) results, as the target mask increases (i.e. illumination area) edge-loss decreases. Gevaux and colleagues (2020) conducted a similar study measuring an opaque light skin coloured sample on a colour checker chart and two translucent materials: skin on the palm of a hand and beige soap. Measurements were taken by three spectrophotometers with varying the illumination/measurement areas: 3.5mm/3.5mm, 4.5mm/4.5mm, 6mm/6mm, 10mm/10mm, 17mm/17mm, and 25mm/25mm, and 20mmx21mm/2mm. For the opaque sample, there was no change when changing illumination area/measurement area. However, the soap and palm of skin found that the when the illumination area/measurement were the same small area (i.e. 3.5-4.5mm) there was more edge-loss causing an underestimating amount of reflectance in comparison to when the illumination area was much larger than measurement area (i.e. 25mm/6mm and 20mm/2mm). This study indicates that when the illumination area is much larger than the measurement area, there is a reduction in edge-loss due to increase of amount of reflectance allotted for measurement. One issue with this study is that the samples are of undefined size and specifically restrict the study to infinite thick objects to prevent light transmission. Gevaux *et al.* (2020) argues that edge-loss occurs when a translucent material is illuminated in a finite area as subsurface scattering creates a region of diminishing light around the point of illumination. However, this claim is based on a study with samples of indefinite size. Teeth have a definitive small size, so what is considered a finite size of illumination? This study suggests the amount of edge-loss is influenced by the relationship between illumination area to measurement area to sample size.

Previous studies have not investigated the amount of edge-loss that occurs when changing the sample size in relation to colour measurement devices. However, studies have investigated how translucency changes

with sample size. Size and level of translucency will affect the amount of edge-loss that occurs as it influences the amount of light travelling through the medium and subsurface scattering. For example, high scattering materials (i.e. opaque materials) have limited subsurface scattering reducing the amount of edge-loss. The degree of translucency varies depending on the depth and width of dentine and enamel layers of the tooth, which can vary between teeth within the same mouth (i.e. canines, incisors, molars), vary within the same tooth (i.e. gingival, central, incisal), as well as between subjects due to age (ASTM E2466-13, 2013; Brook *et al.*, 2007; Yu *et al.,* 2009). For example, the incisal edge is more translucent than in the central location causing more light to pass through the material instead of being reflected back to the detector and may be more effected by background colours (Chu *et al.*, 2010; Joiner, 2004). Enamel thickness was found to be a statistically significant predictor of colour difference. As enamel decreased the colour difference increased, this is due to the translucency of enamel increasing making the dentine colour more prominent (Oguro *et al.,* 2016). Previous studies have evaluated the difference in translucency as size changes. Wang *et al.* (2013) evaluated the difference in translucency parameters of 8 glass and 5 zirconia dental ceramic disks when changing their thickness using a spectrophotometer. All the disks had a diameter of 13mm, but the starting thickness of the glass ceramics were 2.2mm and 1.1mm for the zirconia disks. Results showed that as thickness decreased, translucency increased. However, the zirconia samples, which initially are less translucent than glass, produce a greater difference in transparency values than the glass ceramics. While this study did not directly evaluate edge-loss, it confirmed that as the sample decreases in size more light is able to travel through the sample. These results are further confirmed by Church *et al.* (2017) study in which 4 zirconia samples were compared to a glass ceramic in A2 shade in different thicknesses: 0.5mm, 1.0mm, 1.5mm and 2.0mm. This study found as thickness decreased, translucency increased. One problem with this study is there is no

diameter size of the sample. Would increasing the translucency and size of sample, increase or decrease edge-loss?

Increasing translucency of a sample allows for more light to travel though them, allowing light to exit quicker via the edges and back of the sample compared to less translucent samples producing more edge-loss. This theory could be applied to the width and depth of a sample. A smaller sample would allow more light to exit the sample producing more edge-loss. Arguably, with a small sample more light travelling from outside the measurement area could travel through the material into the measurement area compensating for any edge-loss. However, this is dependent on the illumination area compared to the sample size. If the illumination area was larger than the sample there could be a reduced edge-loss due to light compensation proving the Bolt *et al.* (1994) theory on edge-loss as correct. Would this theory still work when increasing the width and depth of a sample? Thicker and larger samples might have more edge-loss (i.e. lower L* values) as there is less light compensating for light 'lost' from the standard illuminant as it has farther to scatter to the edges and back of the sample. Alternatively, a larger sample might act more like an opaque sample causing less sub-surface scattering producing less overall edge-loss.

With the increased interest in digital cameras for tooth colour analysis, the amount of edge-loss should be evaluated. All previous research in edge-loss uses either a spectrophotometer or a spectroradiometer for analysis. This study investigates the amount of edge-loss that occurs using four different measurement devices: Konica Minolta spectrophotometer, Vita EasyShade spectrophotometer, Konica Minolta spectroradiometer, and a Digital Camera. The amount of edge-loss is evaluated in different sized samples varying in widths and depths. The hypothesis is as translucency decreases the edge-loss will increase. The non-contact based devices will produce less edge-loss than contact-based devices as there is no target

mask restricting light from being measured. As the sample increases in size, the edge-loss will increase as there is more area for light to subsurface scatter through.

4.2 Physical Measurements

According to the American Standard Test Method E1349-06 (2018), translucent materials samples should be measured on a black background in order to prevent backscattering allowing for better evaluation of edge-loss (Johnston *et al.,* 1996; ASTM E1349-06, 2018). A black matte cardstock was used to provide an opaque backing to each sample. All instruments were arranged to have a measurement geometry which ensured the specular component was excluded. Device set up and warm up followed section 2.2.

4.2.1 Samples

A set of 72 cylindrical custom disks were made by Dental Technology Services. The disks were made out of melded block zirconia, a material used frequently in dental restorations due to its appearance and cheap cost (Wang *et al.,* 2013). All samples were cured in the same furnace to improve consistency.

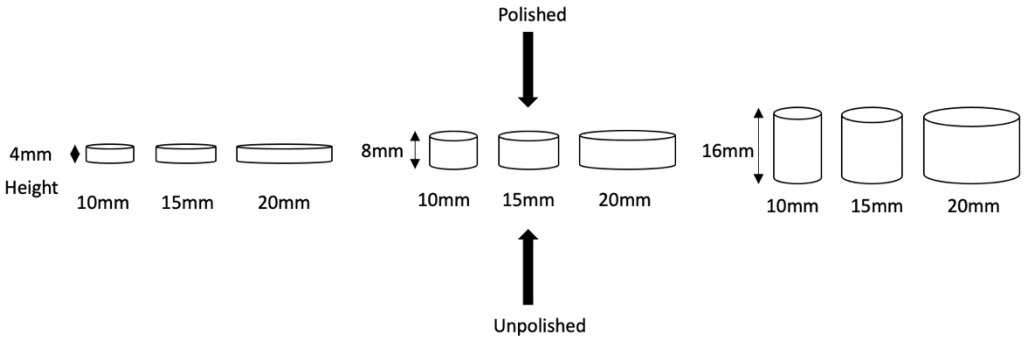

Figure 48: Schematic diagram of the varying sizes of the custom disks

The disks vary in height, width, translucency, colour and finish (Figure 48). One side of each custom disk is unpolished (unfinished) while the other

side of the cylinder is polished (finished). With the two different finishes, there are a total of 144 samples; 72 unfinished and 72 finished samples.

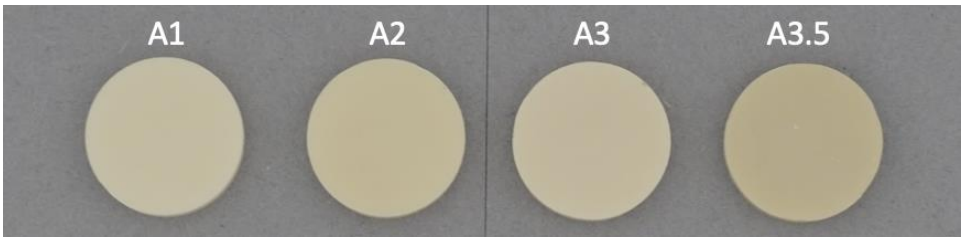

Figure 49: The 4mm x 20mm custom disks displayed in the four shade guide tabs based on Vita Classical Shade Guide

Samples were split into four different sets of 18 cylindrical samples based on colour. The 4 different colour chosen for the samples were on shade guide tabs from the Vita Classic Shade guide set: A1, A2, A3, and A3.5 (Figure 49). The colours were chosen in order to provide a small variety of samples to see if edge-loss is effected by the colour of the sample.

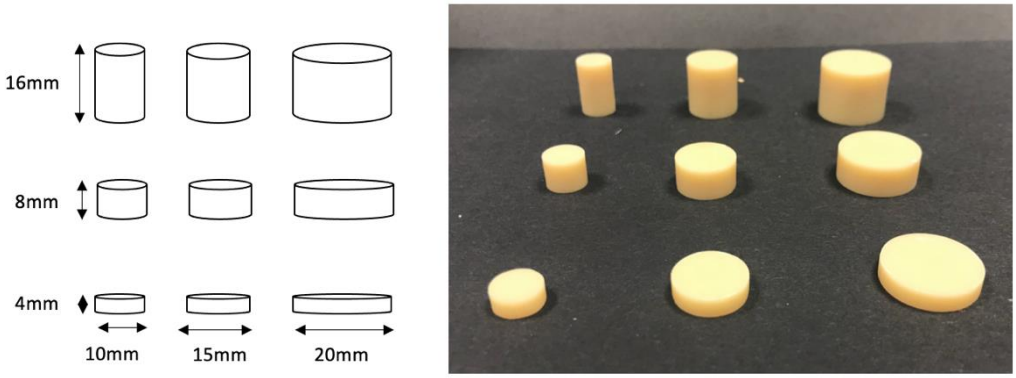

Figure 50: Schematic diagram and photo of the various heights and widths of samples

Each set is split into two different translucencies; 9 highly translucent (1T) and 9 translucent (2T). The different translucencies is to see how translucency effects edge-loss. Each translucent set of 9 are made up of three different widths (10mm, 15mm, and 20mm) and three different heights (4mm, 8mm, 16mm) shown in Figure 50.

4.2.2 Alignment Sheet

The alignment sheets were used for spectroradiometer and DigiEye measurements. An A4 sized sheets were used for alignment as the size

perfectly fits into the tray insert inside the lighting cubes. Three different alignment sheets had to be made for the three different heights of samples to ensure they were equal distance from the capturing lens.

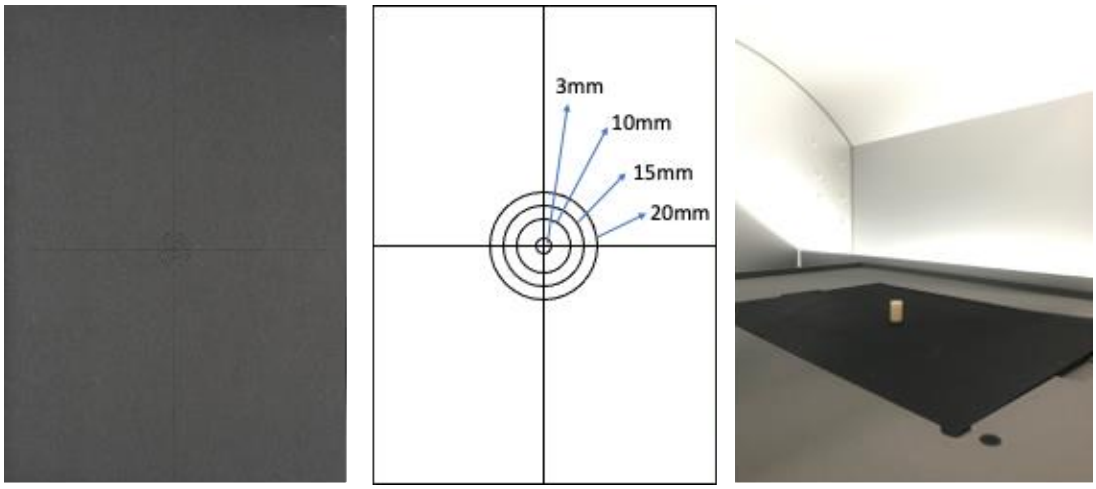

Figure 51: The alignment sheet, schematic diagram for samples with a height of 16 mm, and an image of a sample (10mm x 16mm) on top of alignment sheet

Samples with a height of 16mm were placed in the centre of an A4 sized (210mm x 297mm) black background with four circles marked on it shown in Figure 51. Each alignment sheet had the four circles for the precise central alignment of the different width disks. Going from the outer most circle to the smallest circle is for the different width samples starting with 20mm, then 15mm, and 10mm. The smallest circle is a 3mm diameter to imitate the measurement aperture that of a spectrophotometer, so all devices are roughly measuring the same amount of the samples.

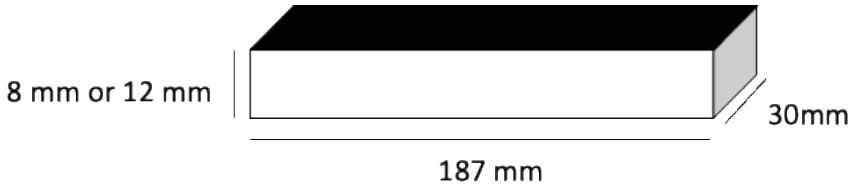

Figure 52: Schematic diagram of custom blocks

In order for the samples to be the same distance away from the measuring device, two custom 3D printed blocks were made for the 4mm and 8mm height samples. One custom block had a height of 12mm for the 4mm

samples and the other block had a height of 8mm for the 8mm samples. Both blocks had a width of 30mm, length of 187mm (Figure 52). Black cardstock with circle temple for each 3D block, 30mm x 187 mm, were placed on top of each block providing a black background for measurements as well as ensuring samples in the same position as all other sized samples for measurement.

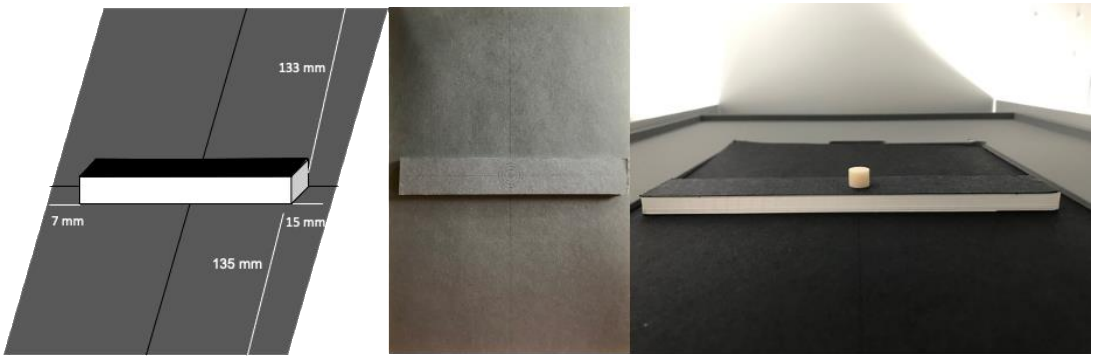

Figure 53: 8mm Alignment Sheet schematic diagram on the left, middle is a photo of the alignment sheet from a birds-eye view, and on the right is the alignment sheet with a 8mm sample sitting on top of the custom block

Each blocks was placed in the centre of a piece of a black cardstock paper. For the 8mm alignment sheet the block was placed 133 mm from the top, 135 mm bottom of the paper, 7 mm from the right and 15 mm from the left (Figure 53).

Figure 54: 4mm Alignment Sheet schematic diagram on the left, middle is a photo of the alignment sheet from a birds-eye view, and on the right is the alignment sheet with a 4mm sample sitting on top of the custom block

For the 4mm alignment sheet the custom block was placed 135 mm from the top of the A4 black cardstock paper, 133 mm from the bottom, 10 mm from the right and 12 mm form the left (Figure 54). The black backgrounds were each alignment sheet were measured three times with a Konica Minolta CM-200d spectrophotometer and averaged to produce respective CIELAB are shown in Table 14.

Table 14 CIELAB values for each alignment sheet's black background

Alignment Template for sample heights	L*	a*	b*
16 mm	26.01	1.12	-1.85
8 mm	25.04	1.03	-1.68
4 mm	25.27	1.13	-1.77

The black backgrounds are slightly different in colour. The colour differences (CIE00) between black backgrounds is less than one unit as shown in Table 15 and will not affect measurements of samples. Average colour difference between the three black backgrounds is 0.41. However, the use of the black background should not influence the colour measurement of each block or effect the amount of edge-loss.

Table 15: Colour Difference between the black backgrounds of the different alignment template sheets

Comparing Black Backgrounds of Various Alignment Sheets	Colour Difference (∆E00)
16mm vs. 8mm	0.49
16mm vs. 4mm	0.55
8mm vs. 4 mm	0.19

4.2.3 Spectroradiometer + DigiEye illumination cube (SR)

For measurements using spectroradiometer an alignment sheets were used in order to ensure that the samples were placed at the exact same 0° geometry location from the colour measurement device for measurement. The custom made alignment sheets were used the top each cylindrical disk was exactly 39.8 cm away from the bottom of the

spectroradiometer lens. The measurement angle of the spectroradiometer was set to 1° providing a measurement area of 6.18mm or a 2.81mm diameter measurement circle. Alignment sheet was placed inside tray of illumination cube, sample was placed in the centre and the door to the cube was closed.

4.2.4 DigiEye + NikonD7000 (DN)

The custom alignment sheets were used so the top of each sample was 50.3cm away from the bottom of the camera lens. The samples were placed at a 0° angle from the lens.

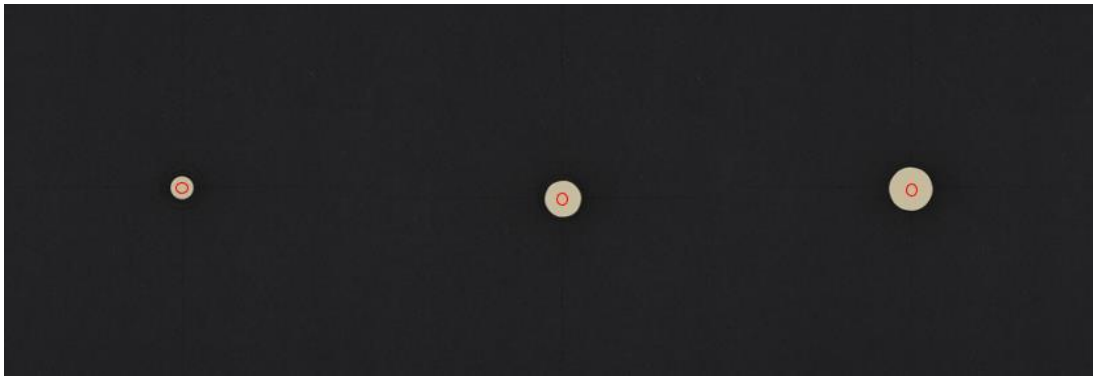

Figure 55: Three different images of samples A1, 1 translucency, height of 16mm with the red circle displaying the 20 pixel measurement area

An image was captured of each sample. The pixel radius of each width 20mm, 15mm and 10mm, was measured and had respective pixel radius' of 70, 110 and 150. A fixed pixel circle radius of 20 was used to measure approximately the centre of each sample which provides a measurement size similar to the 3mm spectrophotometer aperture. Each measurement produces an averaged set of CIEXYZ values that measurement area.

4.2.5 Konica Minolta CM-2600d Spectrophotometer (SPE)

For the custom-disks to have uniform contact with the spectrophotometer, the spectrophotometer was flipped upside down allowing for easier alignment of the aperture with the centre of the disk.

Figure 56: Image of a 10 mm x 16 mm sample with a black piece of paper covering the bottom of the sample for measurement

A black cardstock paper was placed on top of the sample prior to measurement to provide an opaque backing as shown in Figure 56. Specular excluded only measurements were measured.

4.2.6 Vita EasyShade (ES)

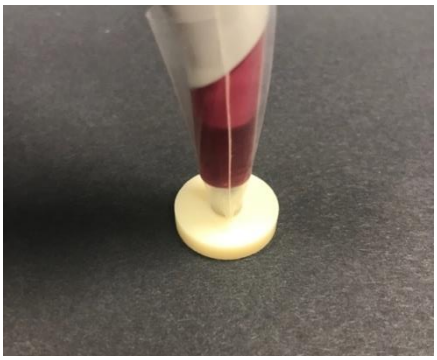

Figure 57: Example of the probe of the EasyShade placed centrally on a 4mm x 20mm sample for measurement

Each of the samples was placed on a black background with CIELAB values of L*=26.01,b*= 1.12,and a*= -1.85. The probe of the EasyShade was placed centrally on each sample as shown in Figure 57.

4.3 Data Analysis

The data for each sample from each instrument were converted into CIELAB values and colour difference (∆ E00) were calculated using the methodology from section 2.6.3, and section 2.7.

4.4 Results

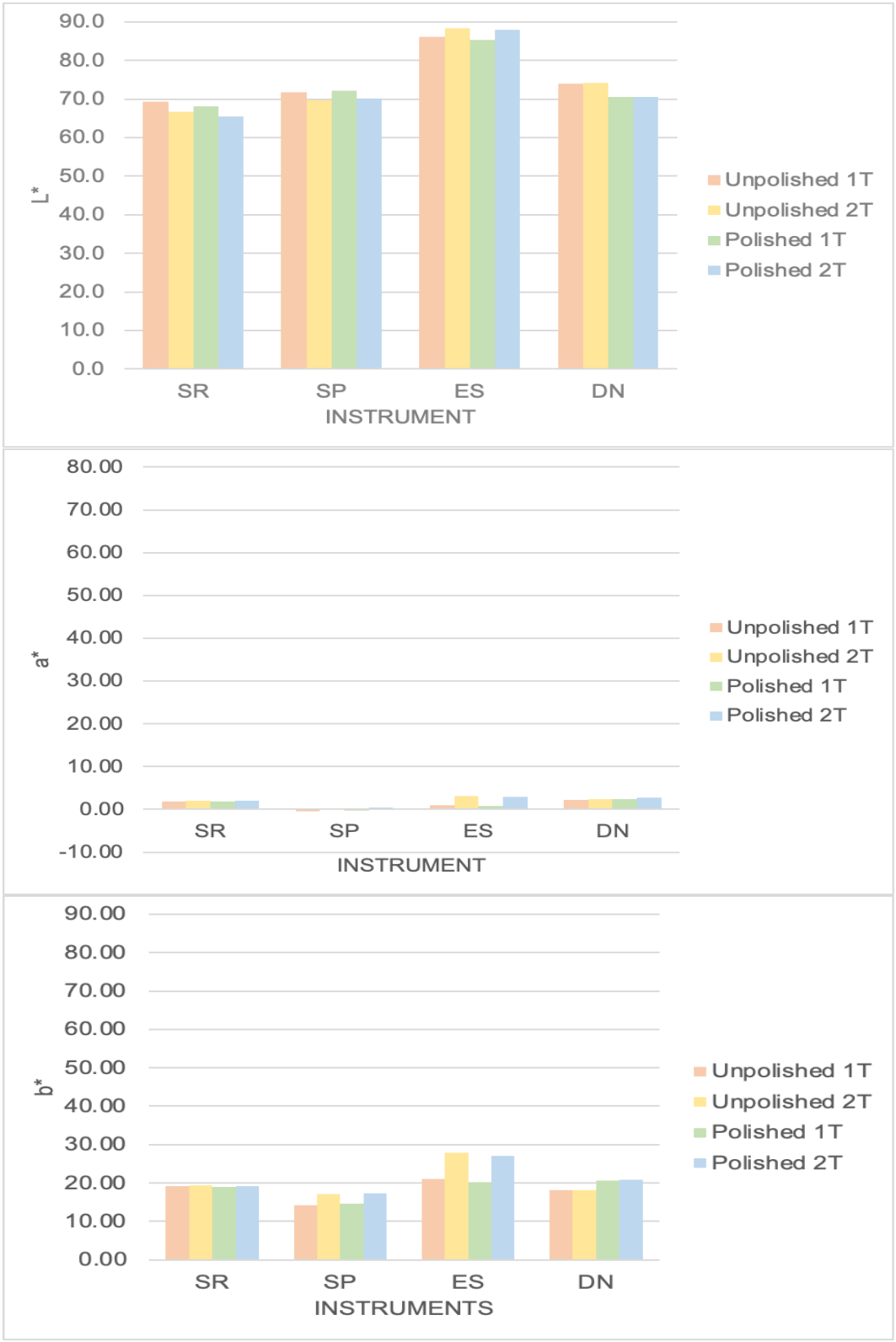

Figure 58: Comparing averaged L*, a* and b* values for their respective instrument for different translucencies and finishes

The CIELAB values were calculated by averaging all the samples of different sizes and shades that were varied in translucencies and finishes for the respective instruments. Figure 58 displays the average L*, a* and b* values for the different translucencies and finishes for their respective instruments. Overall, SR produces the lowest overall L* values (65-69) and ES produces the highest L* values (85-89). The SP and DN produced similar L* values between (70-74). Instruments SR and SP, the 2T sample produced more edge-loss than 1T by a Δ L* between 2 and 2.6. This was the case for both the unfinished and finished samples. ES device produces converse results with 1T producing more edge-loss than 2T for both the unfinished (Δ L=2.3) and finished samples (Δ L=2.6). The DN displayed minimal difference between the two translucencies measurements indicating it is not affected by edge-loss (Δ L=0.1 unfinished and Δ L=0.2 finished). DN is more affected by finish than the rest of the instruments with a difference between finishes of Δ L=4.

In the bottom graph of Figure 58, it can be seen that SP and ES produce larger variation in b* values than SR and DN between translucencies and finished. The SR and DN devices produced the most consistent b* values with Δ b* of 0.1 (SR) and 0.2 (DN) between 1T and 2T, which suggests these devices are less affected by translucency. Instruments SP and ES b* values were affected by translucency with 1T producing lower b* values than 2T with a Δ b*= 7 (ES) and SP with a Δ b*=2.7 between translucencies. Similar to L* results, the translucencies influences the measurement more than the finish for SR, SP and ES. While the finish influences the measurement for DN than translucency factor.

On a comparative scale, the differences in a* may seem minor as the range within the tooth colour gamut is much smaller than L* values.

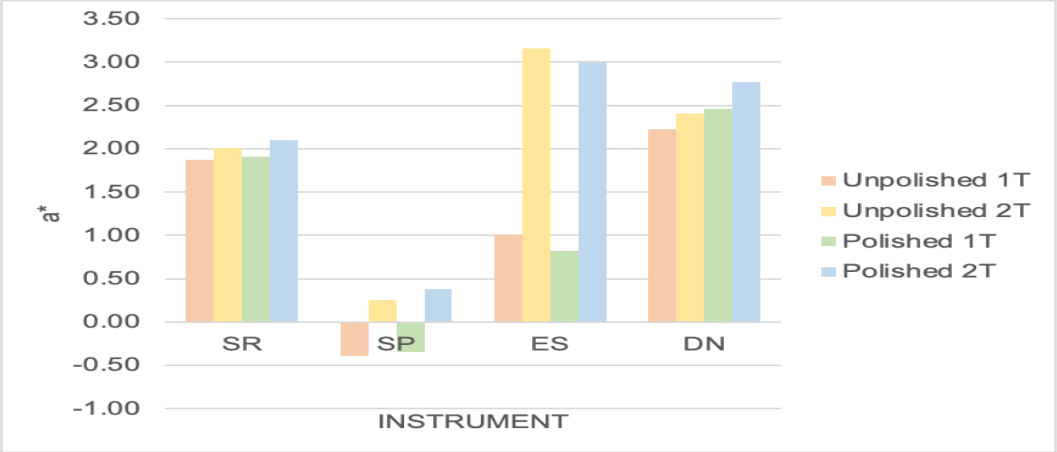

Figure 59: Close up scales of the averaged a* values of the different translucencies and finishes for their respective instruments

Figure 59 displays a close up scale of a* values exhibiting the inconsistency between values due to instrument used, finishes, translucencies. For all the instruments, a lower a* value was produced for the samples which were more translucent. The SP instrument produced interesting results with 1T samples producing a negative a* values, which is a value towards the 'green' hue. While the 2T samples have a positive a* value toward the 'red' hue. The Δ a* between 1T and 2T was 0.6 for unfinished samples and 0.7 for finished samples. The ES produced the largest Δ a* between translucencies with 2T produces 2.2 a* values higher than 1T.

Table 16: The Δ E00 differences between translucency 1T and 2T for the two different finishes for the respective instruments

Instrument	Unfinished	Finished
SR	2.0	2.1
SP	2.4	2.4
ES	4.0	4.1
DN	0.2	0.3

Table 16 display the colour difference Δ E00 between translucencies for the two different finishes for the respective instruments. The ES produces the greatest Δ E00 between translucencies. Overall, DN produces the

most consistent results between translucencies with Δ E00 of 0.2 and 0.3 for the respective finishes: unfinished and finished. This indicates that measurements at the incisal location (most translucent) and gingival location (least translucent) would not be influenced by their translucency when using the DN. SP and SR produce roughly the same Δ E00 between translucencies.

Table 17: The Δ E00 between unfinished and finished finishes for the respective instruments

Instrument	ΔE00
SR	1.5
SP	1.1
ES	1.6
DN	5.3

Table 17 displays the colour difference between unfinished samples and finished samples. SP is the least affected by the finish of the sample. While DN is effected the most by the finish of the sample indicating that the way the teeth are captured is important as teeth are naturally glossy.

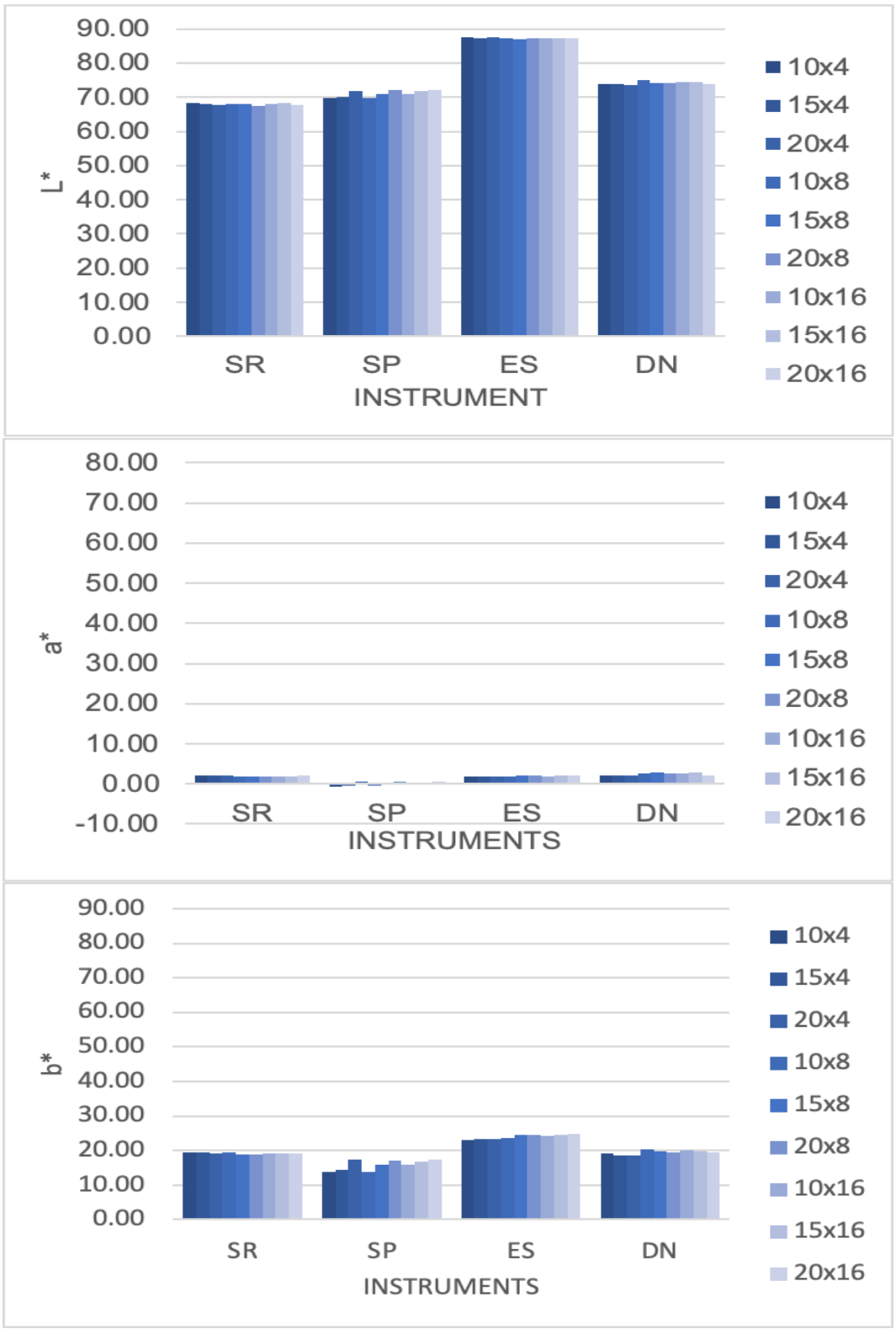

Figure 60: Comparing overall averaged L*, a* and b* values for their respective instrument for the different sized samples

The CIELAB values displayed in Figure 60 are the averaged values of all the samples based on size, regardless of finish and translucency. As the width and depths increase there doesn't seem to be any consistency between devices or sizes. Results indicate that width of the sample influences edge-loss more than the depth. This also varies by sample. For the SR instrument, as the width increases the edge-loss increases. SP produces the opposite results with the edge-loss decreasing as width increases. As the depth increases, the edge-loss decreases. For SR and SP, the depth increases, there is minimal difference in edge-loss. The ES device produced the largest amount of edge-loss for the 15mm width samples, with 10mm with the least amount of edge-loss for both the 4mm and 8mm. At the 16mm depth, the sample with a width of 10mm produced the most edge-loss with samples, with 15mm width having the least amount of edge-loss. For the DN device, the 20mm width produced the most edge-loss. The 10mm width produced the least amount of edge-loss for the 8mm depth samples, but the 15mm produced the least amount of edge-loss for the 4mm depth samples and 16mm samples.

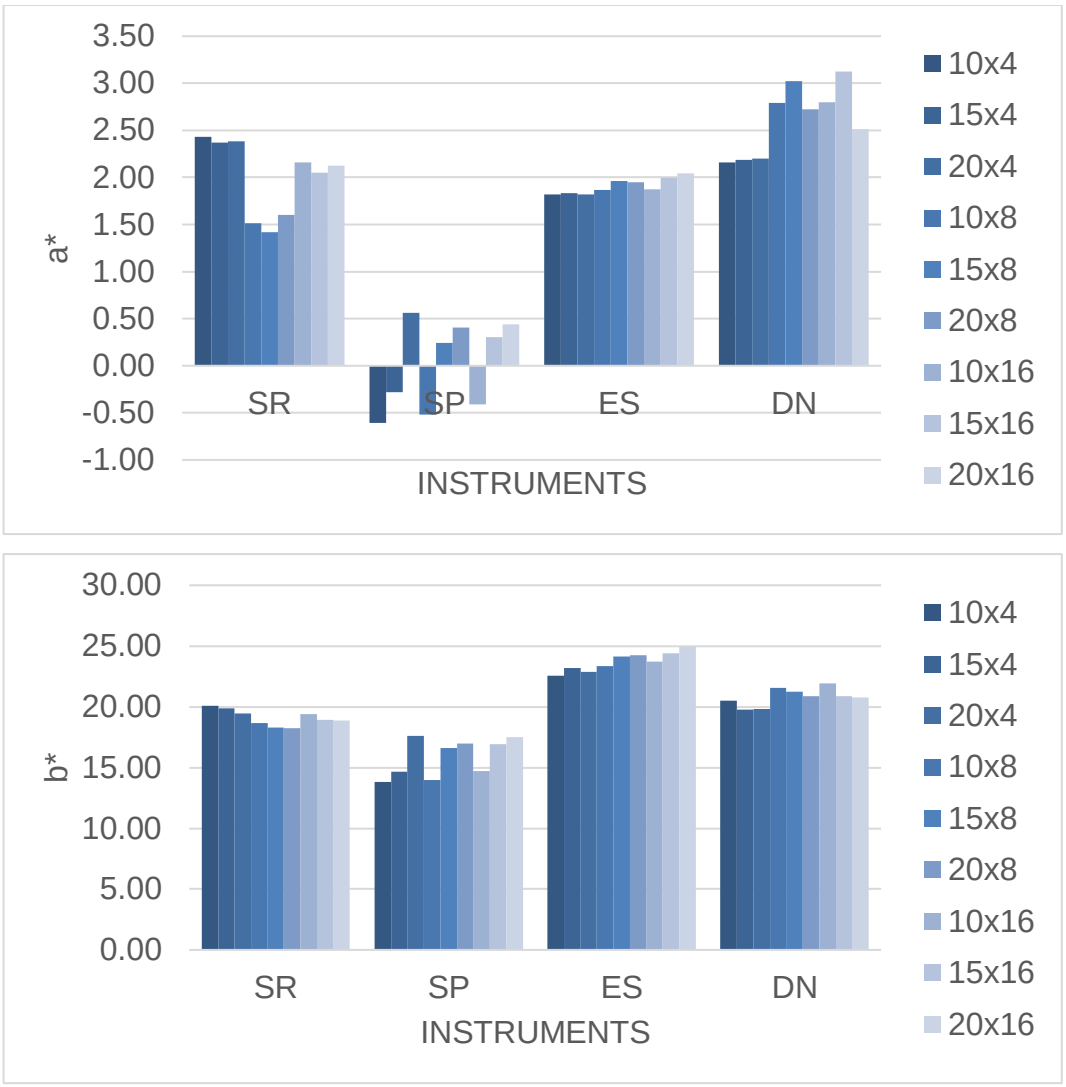

Figure 61: Individual close up scales of a* and b* for their respective instrument for the different sized samples

Figure 61 displays a close up scale of a* and b* comparing the different sized samples for their respective instruments. It can be seen that there are more inconsistencies in a* and b* values for all different sized samples as well as instruments.

Table 18: Differences in CIELAB values and ∆ E00 of all 144 samples between respective instruments

Instrument Comparison	∆L*	∆a*	∆b*	∆E00
SR vs SP	-2.9	2.0	3.5	-3.8
SR vs ES	-19.3	0.0	-4.8	-13.9
SR vs DN	-6.2	-0.5	-0.3	-4.7
SP vs ES	-16.4	-2.0	-8.3	-12.4
SP vs DN	-3.2	-2.5	-3.8	-4.2
ES vs DN	13.1	-0.5	4.5	9.3

Table 18 displays the overall differences in CIELAB values and colour difference between instruments for all 144 samples. A negative value indicates that the first instrument produces a lower values than the second instrument. The ES produced the highest colour values compared to all other instruments. The DN Produced higher overall values than the SP and SR.

4.5 Discussion

Based on the results, the DN produced the most consistent measurements between different translucencies and is the least affected by edge-loss. The ES produced more edge-loss with the more translucent samples following the hypothesis. ES produced the highest L* values, but more importantly the ES measurements produced the most inconsistent a* and b* values across the device indicating the instrument sensitivity to translucency. The SP instrument followed similar pattern to SR in which 2T produced more edge loss than 1T, but the a* and b* produced differing values indicating these characteristics are affected by translucency. The 2T is more opaque than the 1T samples which increases subsurface scattering reducing the amount of light reaching the detector.

All four devices were influenced by the finish of the sample. This is to be expected as matte finishes and glossy finishes have different reflecting properties. The SR, ES and DN produced darker L* values for the finished samples compared to the unfinished samples. This could be because the

measurement geometry is specular excluded. It is interesting that the DN produced the largest colour difference indicating its sensitive to the difference between glossy and unfinished samples compared to other devices. It might be that for tooth colour, the detector should be at an angle versus at a 0° angle.

The SP had higher L* values for the finished samples than unfinished. The measurement port of the SP is at an 8° angle. While set to spectral excluded, having an angled measurement geometry could capture other angled reflected rays. Glossy samples are more saturated when viewed directionally.

It was hypothesized that as the sample increases in size, the edge-loss increases (i.e. producing lower L* values) as the light will be 'lost' within the sample due to an increased area to subsurface scatter through. The results indicate that there is no overall trend when varying the size of the samples. The DN did not produce any specific pattern as the size increased. For the SR, this hypothesis is true. With the entire sample illuminated, more light is able to reach the detector with a smaller sample as there is less area for subsurface scattering than larger samples. In contrast, for the SP as the sample increased in size the edge-loss decreased. A smaller sample allows for light to exit the sample quicker outside the target mask than a larger sample (i.e. width and depth) as there is an increase in area for subsurface scattering which allows for light to be scattered back into the detector. It is interesting that this is not the case for the ES. It seems that the ES produces relatively consistent values as the size of the sample increases, which could be due to the fact that the ES produces spectral values from a database of measurements, selecting the measurements that relate closest to the shade measured.

The hypothesis that non-contact based devices will produce less edge-loss than contact-based devices as there is no target mask to restrict light was not confirmed in this study. This theory was developed based on

results from the studies conducted by Bolt *et al.* (1994) study, Lee *et al.* (2004) and Gevaux *et al.* (2020) which suggest that when the illumination area is much larger than the measurement area there is a reduction in edge-loss (higher L* values), as well as an increase in a* and b* values. In the Lee *et al.* (2004) and Gevaux *et al.* (2020) studies, only spectrophotometers were used varying target masks for comparison. In the Bolt *et al.* (1994) study he compared a SP and SR device, but with no standardised samples. The results from this study indicated no pattern. Both SR and DN were non-contact based devices and the ES and SP both contact based devices. The SR device produced the lowest L* values followed by the SP, DN and ES. The results could be a product of the differences in the amount of light and intensity of the light shining on the samples between devices. The SR device produced lower L* values than the DigiEye even though they are both contactless devices. Both devices use the controlled lighting cube for controlled illumination, so it is interesting that the SR produces darker L* values. The SR device is designed to measures irradiance and the process of converting irradiance data into reflectance data introduces errors. In addition, the illumination cube was designed for use with the DigiEye system. The digital camera is calibrated in order to produce results that are corrected based on the illumination settings produce by the lighting cube. Both the ES and SP are contact based devices shining an intense light onto the sample. The ES produces L* values than the SP device. This could be due to the measurement geometry of the device.

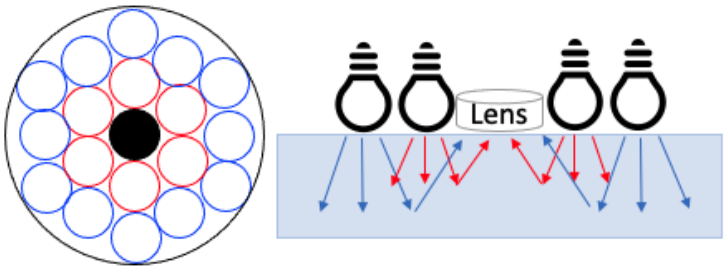

Figure 62: Schematic Diagram of measurement geometry of the EasyShade. On the left is what the probe looks like with two separate light rings (inner in red and outer in blue). On the right is roughly how the light moves.

The EasyShade device does not follow a specific CIE standard measurement geometry. It has been suggested that the EasyShade device produces a pseudocirciular 0°/0° geometry (Paravina and Powers, 2004). It could be argued that the EasyShade produces an angular/0° measurement geometry displayed in Figure 62, which shows a schematic diagram of a rough representation of the EasyShade measurement geometry. The angular geometry (i.e. specular included) could be the reason that the ES produced consistently higher L* values than the other instruments. The other instruments SP, SR and DN all had spectral excluded measurement geometry (i.e. diffuse/0°).

When discussing edge-loss, the main focus is usually differences in L* values. However, the a* and b* values change as well. Unlike the Lee *et al.* (2004) study, the a* and b* results from this study did not produce the same steady increase as the illumination area increased. All the devices produced varying a* and b* values based on changing the translucency and size of the samples. This could influence the overall colour of the tooth producing potentially wrong fillings, crowns, veneers, and inaccurate evaluation of whitening products.

Overall even when measuring the same sample, the different instruments produce different values. This could be due to the different intensities of illumination, amount of illumination, and geometry. While it could be

inferred that the SR is affected the most by edge-loss, followed SP, DN and then the ES. It cannot be said if one instrument produces more edge-loss than another instrument without knowing what the ‘true’ colour values of each of the samples. One instrument might over estimate or under estimate the colour value of the sample. It is unknown which is the most ‘accurate’ instrument to use for tooth colour measurement. This raises the question of how do we define ‘accuracy’ of colour measurement devices for tooth colour measurement.

4.6 Conclusion

All of the instruments produce different CIELAB values for the same samples. SR, SP, and ES are more affected by the degree of translucency than the DN. The DN devices is more influenced by the finish of the sample. The size of the sample affected all of the devices differently. Results determined that both contact-based devices and non-contact based devices are affected by edge-loss. The a* and b* values are equally as influenced as L* by differing translucency, size of sample and device used.

It is unclear if any of the devices produce more edge-loss compared to another device without knowing the ‘true’ value of the sample. It is unknown what the most ‘accurate’ method is for measuring tooth colour. The goal of instrumental colour measurement is produce values which relate to visual perception. This indicates that a new method for defining accuracy of tooth colour measurement should be introduced in which the instrument’s measurements which corelate to visual perception is the most accurate instrument.

5 How to Define the Accuracy of Tooth Colour Measurement

5.1 Introduction

The aim of this chapter is to discuss the concepts of accuracy and precision in the context of tooth colour and then to investigate the most 'accurate' method for measuring tooth colour. Colour is primarily related to the absorption and reflectance properties of materials; but it is also influenced by the angle which light hits an object, angle which the reflected light is detected, and object characteristics (i.e. smooth, rough, glossy, translucent). Colour measurement instruments follow standards set by the CIE. In addition, there are other standard bodies that have published standards, advice and guidance about the colour measurement of opaque materials (ASTM E1347-06, 2020). However, only one standard that is specific to tooth colour measurement exists.

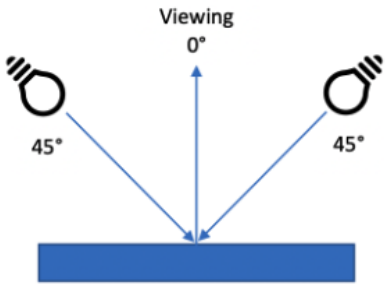

Figure 63: Schematic diagram of measurement geometry used for tooth measurement according to ASTM E2466-13

ASTM E2466-13 (2013) describes a standardised method for measuring tooth colour using a digital camera and specifies a 45°/0° geometry with two light sources at 45° angle and 0° viewing angle (Figure 63). While this standard has recently been withdrawn, this geometry has been used in other dental research with either spectroradiometers, spectrophotometers, colorimeters, or a high resolution digital cameras (Paravina *et al.*, 2021, Guan *et al.*, 2005; Gozalo-Diaz *et al.*, 2007; Paul *et al.*, 2002; Cho *et al.*, 2007). Luo and colleagues (2009), as well as Pérez and colleagues (2016), used a spectroradiometer at a 45° measurement angle with a

lighting cabinet providing approximately 0° illumination. Others have used a diffuse/0° geometry taking measurements with a hemispherical spectrophotometer (Lim *et al.*, 2010; Paravina *et al.*, 2021; Seghi, 1990). In a clinical trial, an illumination ring was placed around the lens of a camera (0° measurement angle) (Luo *et al.*, 2007). Other researchers have used different devices such as dental-specific spectrophotometers such as Vita EasyShade or Shade pilot; others have set up different measurement geometries which do not follow CIE standards (Lim *et al.*, 2010; Lasserre *et al.* 2011; Kim *et al.*, 2018; Parameswaran *et al.* 2016; GJohmez-Polo *et al.*, 2014; Elamin *et al.*, 2015). Even with the ASTM standardised test method, instrument and measurement geometry remain inconsistent in dental research. The evidence, however, that the ASTM method is the most accurate method to measure tooth colour is limited. With numerous devices available, it is uncertain which method will measure a tooth's true colour. But what is the true colour of a tooth and how do we define accuracy in dentistry?

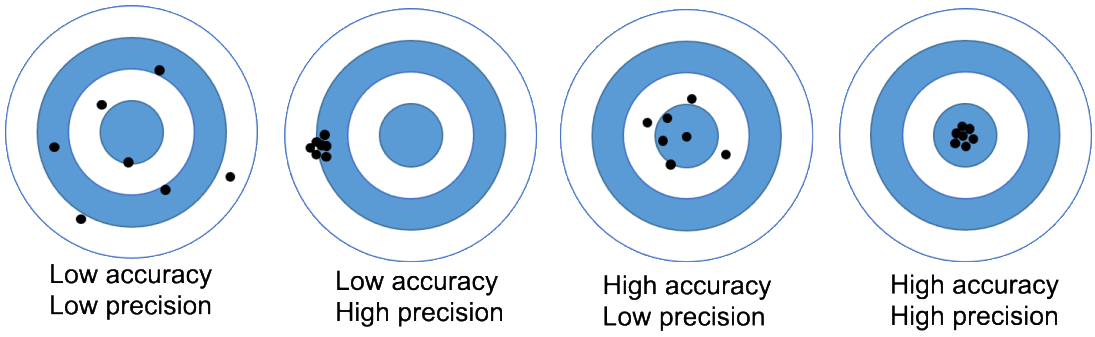

Figure 64: Visual representation of the between accuracy and precision based on a target

Accuracy is the ability of a measurement to match a 'correct'/set value, while precision is defined as the repeatability of a measurement (Nedelcu *et al.*, 2018; Johnston, 2009). Figure 64 visually illustrates the concepts of accuracy and precision; in the figure, the centre of the target is considered to be the set or correct value. An ideal method for tooth colour measurement would have high accuracy and high precision and would

deliver repeat measurements that consistently match the correct value. However, whereas it is trivial to measure precision (it is simply the variability of repeat measurements) it is more challenging to measure accuracy since we need to answer the question of what the true or correct value is.

It is clear that the true value must, to some extent, be determined visually. Some researchers have defined accuracy as an instrument's ability to determine the correct shade guide tab during restoration or simulated restoration (Lehmann *et al.*, 2010; Wee *et al.*, 2006). Kim-Pusateri *et al.* (2009) evaluated the accuracy of three dental spectrophotometers and one digital camera spectrophotometer in their ability to correctly measure and identify a known shade tab and found the VITA EasyShade correctly identified the tab in 92.6% of cases (ShadeVision 84.8%, SpectroShade 80.2%, ShadeScan 66.8%). This approach evaluated the accuracy and precision of an instrument's ability to identify shade guide tabs and this was possible because these specialised devices are able to indicate (presumably through an internal database) which shade guide tab the measurement most closely matches. In a study conducted by Paul *et al.* (2002) visual assessments (to determine the closest shade guide tab) of one upper central incisor from 30 patients were made by three dentists and measured with a spectrophotometer by each dentist. Only 26.6% of the time did all three dentists agree on the same tab, while the three spectrophotometric measurements agreed 83.3% of the time. The study concluded that spectrophotometers offered a 33% increase in 'accuracy' compared to visual assessment. However, the data really indicated that the instrument was more *precise* than the visual system and did not measure accuracy. In this case, we have to either accept that the visual system is the gold standard (in which case, by definition, the instrument cannot be more accurate) or, if we really want to compare the accuracy of the instrument and the visual system, we require a third evaluation that is considered as the 'gold standard' against which both systems could be

compared. Bahannan (2014) found that 204 students were able to use the EasyShade to identify a shade tab 80% of the time versus visually they could only identify it 36% of the time. In this case, the identity of the shade guide tab was known to the experimenters and therefore there was objective 'true' value; the experiment therefore indicated that the instrument was more accurate than the visual system. However, again, in this context it is not really about colour measurement per se since no matter how accurate or inaccurate the colour measurement is, an appropriate look-up table could be used to identify the correct tab. In the study conducted by Da Silva *et al.* (2008), visual assessment was used explicitly to ascertain accuracy; a crown made from a shade chosen by visual assessment and a crown made from a shade identified by a spectrophotometer of the incisor of 36 participants was compared to the respective incisor. Three examiners visually rated each crown (1-10 with >8 accepted and <7 rejected) and the mean was chosen as accepted for the participant. The study found that digital spectrophotometers were classed as acceptable for prosthetic dental matching with 78% accepted crowns versus only 22% accepted crowns made from visual assessment (Da Silva *et al.*, 2008). Khurana *et al.* (2007) explicitly assessed precision for three instruments and found that the precision of the Spectroshade Micro (an imaging spectrophotometer) was better than that of an X-Rite ShadeVision (a colorimeter) and a Vita EasyShade (a spot spectrophotometer). The papers reviewed in this section so far are based on two definitions of accuracy: (1) where the identity of the shade guide tab is known and the task of the instrument is to correctly identify it and (2) where the visual system is used to evaluate, for example, the appropriateness of a crown. The first method is objective whereas the second is more subjective, using the visual system as the final arbiter. However, in all cases the criterion for success is based on a shade-tab determination. For the assessment of accuracy, a reference instrument is required. Shade guide determination is not useful if the 'goal' is to evaluate the instrument's suitability for measuring whiteness and yellowness.

Alternatively, accuracy has been determined by comparing a test instrument to a reference instrument (Johnston, 2009; Lehmann *et al.*, 2010; Paravina *et al.*, 2021). In many industries, reflectance spectrophotometers are considered to be the most reliable method for colour measurement against which the accuracy of other instruments (i.e. colorimeters, spectroradiometers, and digital cameras) are compared. While this is a sensible approach for flat opaque materials (i.e. paints, textiles), it is unclear whether spectrophotometers should be used as the gold-standard for in-vivo measurements. The translucency, curvature, structure, size, translucency, polychromaticity, and gloss of teeth increases the complexity of accurate colour measurement. Teeth are specifically prone to edge-loss due to their translucency; a phenomenon where light enters the tooth, is scattered and exits the edge or back of the tooth thus not being measured by the instrument and resulting in darker colour measurements than would otherwise occur (Ahn and Lee, 2008; Johnston *et al.*, 1996; Bolt *et al.*, 1994). The amount of edge-loss is dependent on instrument attributes such as illumination area, contact versus non-contact, measurement geometry, and aperture size (Johnston *et al.*, 1996). In 1994, Bolt *et al.* compared measurements from a spectroradiometer to those from a spectrophotometer with different apertures (3mm, 4mm, and 5mm). They found that edge-loss was wavelength dependent and resulted in a reduction in average L* values from about 75 to less than 55, a reduction in average b* from 17 to less than 4, with much smaller changes to a*. The changes were also found to depend upon the size of the measurement window; edge-loss increased as the measurement aperture decreased in size (Bolt *et al.*, 1994). Instrument manufacturers attempt to minimise edge-loss by illuminating a wider area of the sample than the measurement area (ASTM E1164-12(2017), 2017). However, even with this adjustment, spectrophotometers are still prone to edge-loss. The phenomenon of edge-loss in particular must raise a question about whether contact-based spectrophotometers

should be used as gold-standard reference instruments for tooth measurement. Indeed, it raises the wider question of what the 'true' reflectance measurement of a tooth sample is.

Despite these questions, spectrophotometers are still sometimes used as reference instruments to evaluate tooth colour for 'whitening' treatments, for the evaluation of tooth discoloration or aging, and for identification of shades for tooth restorations (Pop-Ciutrila *et al.*, 2016; Beltrami *et al.*, 2014; Ahn and Lee, 2008; Brook *et al.*, 2007; Oliveira *et al.*, 2015; De Bragança *et al.*, 2021; Marques Junior *et al.*, 2021). It has been suggested that spectrophotometers are the most common and widely used method for tooth colour measurement due to their sensitivity, 'accuracy', and reproducibility (Lasseree *et al.*, 2011; Da Silva *et al.*, 2008). A study comparing two dental spectrophotometers measured the CIELAB values of 102 participants (Khashayar *et al.*, 2012). The correlation between the CIELAB values from the two instruments was extremely poor (r^2= 0.51 between the L* values). In a more recent study, 50 students measured an extracted tooth with two different spectrophotometers (Blum *et al.*, 2018). The CIELAB values from the two spectrophotometers were vastly different from one spectrophotometer (on average L*=0.46-3.04, a*=1.14-1.45, and b* =1.72-2.14) compared to the other spectrophotometer (L*=0.48-1.90, a*=0.05-0.26, and b* =0.29-2.52). This signifies that even two instruments of the same type can have varying CIELAB measurements.

Other studies have used colorimeters as reference instruments. Similar to spectrophotometers, two different colorimeters can also produce different measurements. Cho *et al.* (2007) measured 564 teeth (maxillary and mandibular anterior teeth) with two different colorimeters and found that one colorimeter produced darker, greener and bluer measurements (L* =39.0-65.8, a*=-5.1-4.0 and b*=-1.0-15.1). The other colorimeter (Shade Vision) produced higher L* values (64.5-83.2) and higher a* (a*=1.6-9.8) and b* (b* =10.4-29.0) readings. Lehmann *et al.* (2010) compared three in-vivo spectrophotometers (Vita EasyShade Advanced, Vita EasyShade

Compact, and DeguDent Shadepilot) and an in-vivo colorimeter (Shade Vision) to a benchtop reference spectrophotometer. Two of these devices are no longer available on the market, but the EasyShade devices (advance and compact) were determined to be not as 'accurate' as the spectrophotometer reference device. The spectrophotometers underestimated the L* values compared to the imaging colorimeter which consistently produced higher L* values. This is consistent with the idea that edge-loss occurs with contact-based instruments. However, it does not address which instrument is 'correct'.

Some suggest that colorimeters are less 'accurate' because instead of measuring spectral data they directly measure colorimetric values (CIE XYZ) using broadband filters (Chu *et al.,* 2010; Joiner and Luo, 2017). However, it is not clear that this should make them inherently less accurate that spectrophotometers; besides, colorimeters are themselves prone to edge-loss since they are also contact-based instruments. In clinical situations, it has been suggested that non-contact based systems are preferred because they reduce edge-loss and stray light (Guan *et al.*, 2005; Bolt *et al.*, 1994). Should a spectroradiometer be used instead as the reference standard to which all other devices could be compared?

Spectroradiometers do not present the same challenges as spectrophotometers for measuring dental materials. Since they are non-contact based device they do not have to worry about *in vivo* issues such as contamination or having the aperture make full contact with the curved material (Bolt *et al.*1994). Some edge loss could still occur but this would likely be much less than for a contact-based spectrophotometer. Measurements from a spectroradiometer are usually taken from a distance of at least a metre and require an external light source. The instrument is designed to measured radiance reflected from an object (Joiner and Luo, 2017) or light emitted from a light source. In order to convert spectral radiance to spectral reflectance factors it is necessary to measure the spectral properties of the external light sources; this is

achieved by measuring a white tile (which is assumed to perfectly reflect the light source). The spectral radiance of a sample is divided by the spectral radiance of the white tile at each wavelength to provide the reflectance factor. However, in practice white tiles do not have perfect reflectance and therefore the method, if not corrected, might underestimate the intensity of light source and hence overestimate the reflectance factors. A correction can be applied by measuring the spectral reflectance of the white tile using a reflectance spectrophotometer. However, the tile's reflectance factors may depend upon the geometry of the measurement and this can result in systematic errors if the measurement geometry of the spectrophotometer does not perfectly match that of the spectroradiometer. It is suggested that when spectroradiometers measure opaque materials the measurement geometry should be 45°/0°, while many spectrophotometers used diffuse/0°. In addition, variations in ambient illumination can affect the measurement from a spectroradiometer. The challenges to obtain accurate measurements of spectral reflectance explain why spectrophotometers are more widely used in many industries and used as the reference instrument. Both devices present a number of challenges for producing accurate colour measurement of dental materials making it unclear which instrument should be considered the reference standard.

With advancements in photography and digital technologies there is increasing use of digital cameras. There is growing interest in using digital cameras for the assessment of tooth colour as they are also able to capture texture, shape, perceived translucency, and allow for selection of colour measurement aperture (Wee *et al.*, 2006; Jarad *et al.*, 2005). Brandt *et al.* (2017) measured the incisors of 107 participants and found that the 3Trio inter-oral scanner was 43.9% accurate to the reference EasyShade spectrophotometer. He and colleagues (2020) compared a digital camera with and without cross-polarising filters with a spectrophotometer (ShadePilot) for 50 extracted human maxillary incisors. The polarising filter

removes specular light. The correlation between CIELAB values from the digital camera with no filter and the spectrophotometer were high (r^2=0.93 for L*, r^2=0.92 for a*, and r^2=0.82 for b*). It is interesting to note that the b* values had less correlation. The cross polarised (specular excluded) digital camera measurements produced weaker correlation with the spectrophotometer (r^2=0.88 for L*, r^2=0.0.95 for a* values, and r^2=0.84 for b*). However, the study did not provide the CIELAB values themselves. A similar study by Mahn *et al.* (2021) compared a digital camera with polarising filter with a Vita EasyShade spectrophotometer of 60 maxillary right central incisors. The digital camera with filter produced lower L* values with a larger range (from 52 to 92) compared to the EasyShade L* values (from 69 to 96). The EasyShade had a larger range of a* values (-6 to 2) compared to the digital camera a* values (0 to 7). The two instruments recorded similar b* values. In a study conducted in 2005 by Jarad *et al.* compared CIELAB values from a digital camera with those from a spectrophotometer for 16 shade guide tabs and found that the digital camera had higher L* values (58 to 74) than the spectrophotometer L* (48 to 59). In a more recent study, CIELAB values from an Apple iPhone 8 cell phone were higher than those from a reference spectrophotometer (De Bragança *et al.*, 2021). The precision and accuracy of these systems is influenced by the quality of camera, proper calibration, imaging process model, and lighting conditions (Chu *et al.*, 2010; Tung *et al.*, 2002; De Bragança *et al.,* 2021).

A further underlying problem with most imaging systems is that rather than record XYZ like a colorimeter, they record red, green, blue (RGB) values at each pixel location. Methods are required to convert these RGB values into colorimetric values and this is not trivial. A camera must satisfy the Luther condition to be considered colorimetric, which is when a camera's spectral sensitivities (RGB channels) are the same as the CIE XYZ colour matching function (or a linear transform of them) (Finlayson and Zhu, 2019). There is no such camera on the market and therefore mathematical

approaches are used to approximate CIE XYZ from RGB values. Corresponding RGB and CIE XYZ values are obtained during the calibration process with the use of a colour calibration chart containing a large number and range of colours and mathematical transform is optimised to convert image RGB values to approximate CIE XZY values. The nature of the colour chart used could influence the accuracy of the mathematical conversion; this raises the question of which colour chart to use, how many colour patches should be required on a colour chart, and how should they be distributed throughout colour space? Should there be a specific chart designated for dental materials? If so, then how will the CIE XZY values of those patches be measured? It is not obvious whether to use a spectroradiometer, spectrophotometer or colorimeter (leading to a circular problem). In many studies the set-up of digital cameras is unique and replicating the set-up in a different laboratory would be difficult; variables include the light sources (i.e. ring around camera lens, two or more light sources), digital camera choice, measurement geometry and software, colour chart etc. There is a commercially available system called DigiEye made by VeriVide (2020) which is a colour measurement system that uses an enclosed environment with a digital camera allowing for different uses and research groups to obtain identical results. While this is suitable for in vitro measurements for research and development departments, it is not suitable for in vivo measurements of teeth (i.e. restorations, clinical trials, or evaluating in-office whitening treatments).

In summary, while it is relatively straight forward to assess precision, it is challenging to assess accuracy for an instrument that is designed to measure tooth colour. It is unclear whether reflectance spectrophotometer or spectroradiometers should be considered as reference instruments. It is far from certain whether the notion of ‘true’ spectral or colorimetric values for teeth even make sense. Regardless, it is a fundamental basis of colour measurement that the visual system should be the final judge. In this chapter the accuracy of various instruments is assessed by the extent

to which the measured data correlate to visual perception. More explicitly, whiteness and yellowness indices are calculated from the instrument data and the correlation of these indices with corresponding visual scale data (visual assessment of whiteness and yellowness) is used as a measure of accuracy. The greater the correlation between the instrument and the visual system, the greater the accuracy of the instrument.

5.2 Experimental

In this study, measurements of tooth colour dental materials are made with two spectrophotometers with different measurement geometries, one dental specific spectrophotometer (EasyShade), one spectroradiometer with two different measurement geometries, and one digital camera. Differences between samples are evaluated using whiteness and yellowness indices commonly used in dentistry. These devices are compared to the evaluation of dental materials by a panel of observers, the ratings given by the observers answers are considered to be the 'reference standard'.

5.2.1 Samples

Vitapan shade guide tabs are widely used in dentistry for specification of shades for tooth restorations and bleaching treatments. Due to their curved surface, shade guide tabs are not ideal for this experiment as the curvature could induce measurement errors. In order to ensure complete contact, for contact-based devices, an additional set of samples was used; these were flat custom-made shade tab disks fabricated out of dental porcelain by Vita (Vita VMK Master, Vita Zahnfabrik, Bad Sackingen, German). This material is widely used in dentistry due to its realistic aesthetic properties (Beltrami *et al.*, 2014).

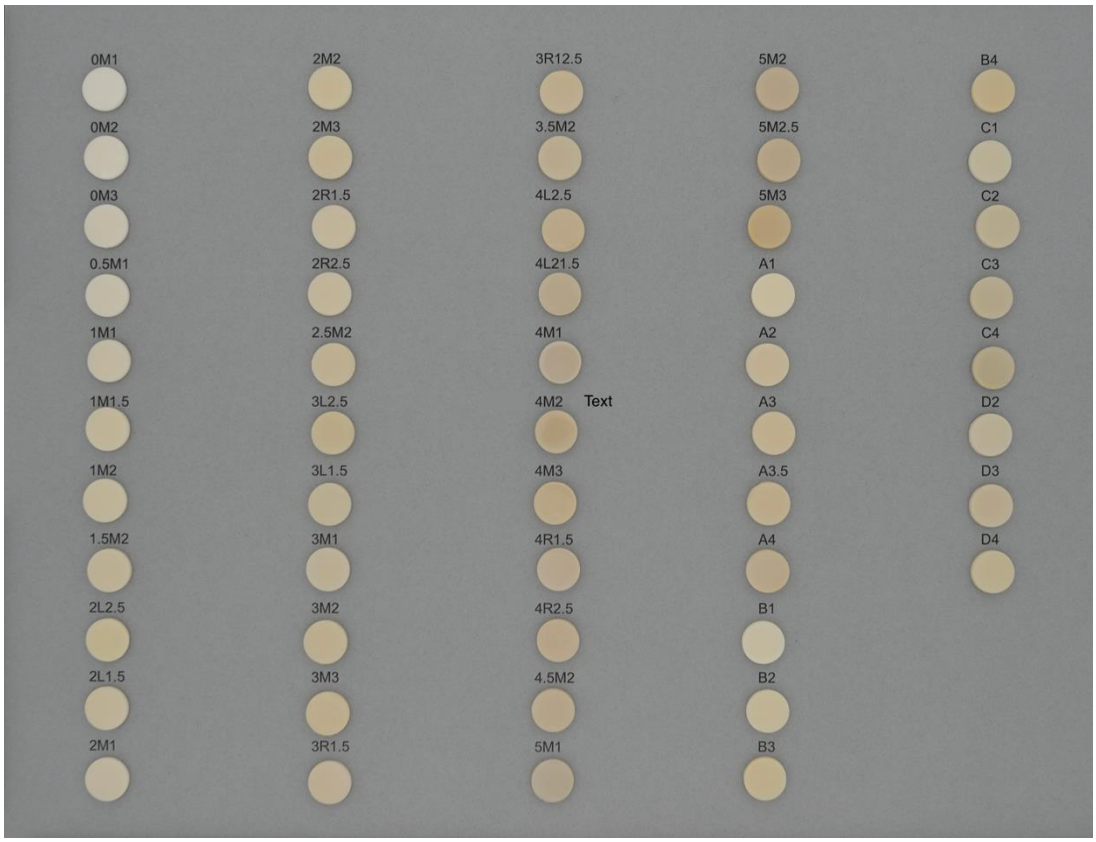

Figure 65: The 52 flat custom ceramic dental disks with their respective Shade Guide Tab names

These samples were used in a previous study conducted by Pan *et al.* (2018). The 52 custom disks were designed to cover the gamut of tooth colour and each was made to be a close match to one of the tabs from the Vita bleachedguide 3D Master, the Vita Classical, or the Vita toothguide Master (Figure 65). Normally the three shade systems have a total of 59 tabs, but some shades overlap in Vita bleachedguide 3D Master and Vita toothguide Master which is why only 52 samples were made.

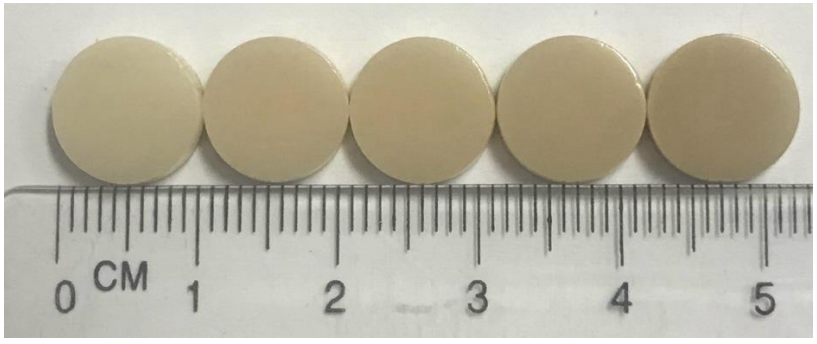

Figure 66: A small selection of the 52 custom disks to illustrate their size.

Shade guides are used to help with dental restorations. One type of dental restoration procedure comprises three layers using different shades with variable thicknesses to obtain a natural tooth appearance (Beltrami *et al.*, 2014). The three layers include an opaque structural layer (typically a metal support with an opaquing agent covering the surface or just an opaque substructure), an intermediate uncoloured translucent layer and an outer shade layer (Blair, 1985). The custom disks developed follow this design. Each of the 52 were cut to be 10.4 mm in diameter with a total thickness of 3 mm (Figure 66). The average width of a maxillary central incisor ranges from 7 mm to 10 mm, maxillary lateral incisors from 5.5 mm to 8 mm, maxillary canines from 6.5 mm to 9 mm, mandibular anterior incisors ranged between 4.5 mm to 6mm, mandibular lateral incisors between 4.5 mm and 7 mm, and mandibular canines between 5.5 mm and 8 mm (Chu, 2007; Chu and Okubo, 2008; German *et al.*, 2016). While the maximum width for a tooth is 10mm, the diameter of these samples was chosen to be 10.4 mm in order for instruments to be more easily measure the samples. The thickness is comprised of three layers; a 0.7 mm layer to represent enamel at the top, 1.7mm layer in the middle for dentin and an opaquing base layer of 0.6mm representing the opaquing layer used in restorations. It is used as a method of masking the colour of the metal structure that is sometimes used at the core of a crown or bridge, mask the metal post that is fitted in root of a tooth or implant to the support the crown, or to make the underlying colour of the tooth. The base layer has a depth of 0.6 mm which is a common depth for an opaquing layer in metal-ceramic crowns (Ahmad, 2006).

5.2.2 Psychophysical Experiment

Previously published data by Pan, Westland, and Ellwood (2018) was available and re-used for evaluation in this study. A total of 80 individuals, pre-screened for colour-blindness using an Ishihara Test, participated in a psychophysical experiment. The experiment was conducted in two different geographical locations with 40 individuals from Leeds (United

Kingdom) and 40 participants from Beijing (China). In each group of 40, there were 10 young females, 10 old females, 10 young males, and 20 old males. Young was classified as any individual between the ages of 18-30, while old was between the ages of 30 and 60. Each participant was asked to rank the 52 disks in order of whiteness, from most white to least white. Individuals were asked to rank the samples in a viewing cabinet with D65 lighting against a neutral grey card, which was provided for each viewing cabinet in the two different geographical locations so the same background was used in study. Samples were observed at approximately 45°.

During the same experiment by Pan et al. (2018), the same pre-screened 40 participants in the United Kingdom were asked to rank the same 52 circular disks based on perceptual yellowness under the same conditions. The yellowness data were not previously published but were available for analysis in this thesis.

5.2.3 Physical Measurements

Each of the 52 disks were measured three times by a spectrophotometer, 45°/0° spectrophotometer, spectroradiometer, a Vita EasyShade, and a DigiEye system. The spectroradiometer was used twice in two different measurement geometries. Table 19 describes instrument conditions including abbreviation, measurement geometry, illumination area, measurement area and if the device makes physical contact with the sample or not.

Table 19: Description of conditions of each instrument

Instrument	Abbreviation	Illumination Angle/ Measurement Angle	Illumination Area	Measurement Area	Contact vs. Non-Contact
Spectroradiometer	SR	Diffuse/0°	≈ all	2.83mm	Non
Spectroradiometer	SR45	45°/0°	≈ all	3.04 mm	Non
Spectrophotometer Specular component included	SPI	Diffuse/8'°	6mm	3mm	Contact
Spectrophotometer Specular component excluded	SPE	Diffuse/0°	6mm	3mm	Contact
Spectrophotometer	SP45	0°/45°	6mm	4mm	Contact
Vita EasyShade	ES	Unknown	Unknown	1mm	Contact
DigiEye	DN	Diffuse/0°	≈ all	≈ 3mm	Non

5.2.3.1 Spectrophotometer

5.2.3.1.1 Konica Minolta CM-2600d - Diffuse/0°

Figure 67: A custom dental disk placed central on 6mm/3mm target mask of reflectance spectrophotometer

For the custom-disks to have uniform contact with the SP, the SP was flipped upside down, allowing easier alignment of the aperture with the centre of the disk (Figure 67). This instrument is safe to be used upside down as long as you ensure no dust gets into the chamber (Konica Minolta, 2014). The specular component included (SPI) and specular component excluded (SPE) data for each sample were separately analysed and converted into colorimetric data.

5.2.3.1.2 X-rite 962 - 0°/45°

Figure 68: Custom disk placed central on target mask

The SP45 was placed up-side down in order for the custom disk to have uniform contact with the target mask and ensure the disk was placed central over the aperture. This instrument has a ‘shoe’ in which the target mask locks into. The custom disk is placed in the centre of the target mask (Figure 68).

Figure 69: Images of before and after of instrument shoe being pressed into instrument body in order to trigger measurement collection of one of the custom disks

With the custom disk on top of the target mask, the instrument shoe is pressed down gently in order to prevent the custom disk from moving from its placement (Figure 69).

5.2.3.2 Spectroradiometer

5.2.3.2.1 Diffuse/0°

Figure 70: Konica Minolta CS-2000s tele-spectroradiometer arranged on a VeriVide DigiEye lighting box to provide a measuring geometry of diffuse:0°

The spectroradiometer was configured on top of a VeriVide DigiEye lighting cube in order to obtain approximate diffuse/0° geometry. The SR was positioned at the 0° angle above the sample with the VeriVide DigiEye lighting box providing roughly diffuse lighting (Figure 70). The focal length was set at roughly 0.42. The measurement angle was set to 1°. With a 50 cm distance from object lens to sample with a 1° measurement angle, the aperture area is 7.78mm. The distance between the bottom of the objective lens to the top of the sample was 41cm, which produces a measurement area of 6.37mm or aperture diameter of 2.83mm.

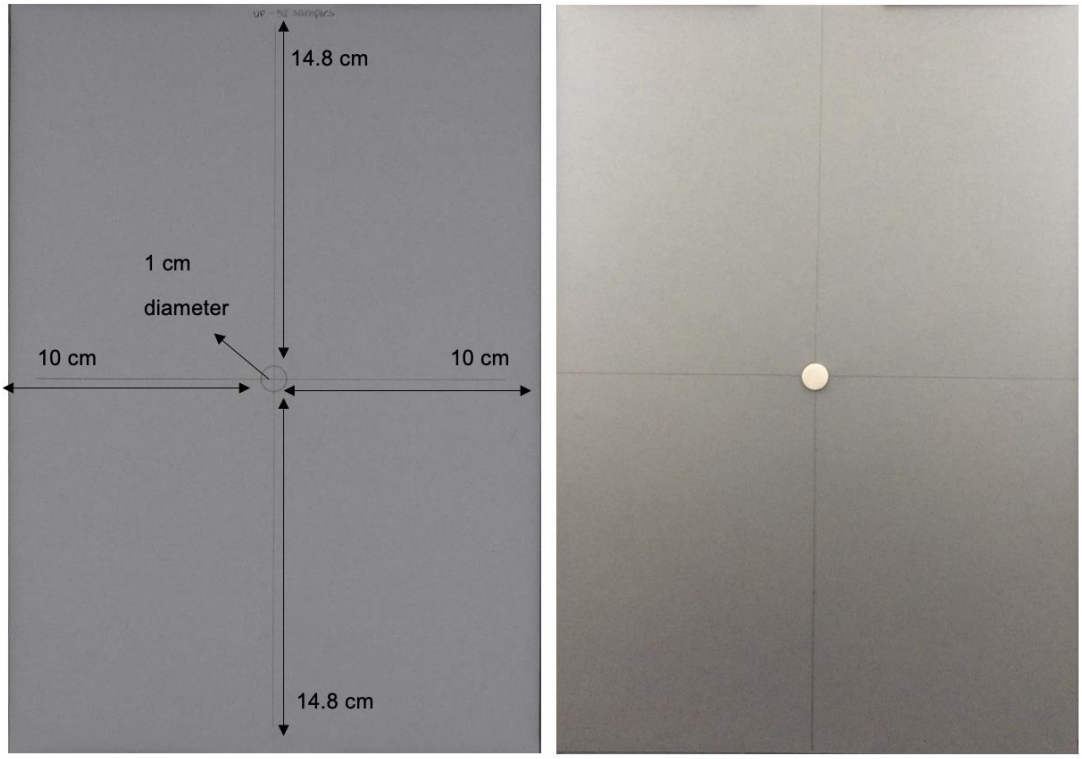

Figure 71: On the left is the A4 grey cardstock alignment sheet for custom 52 disks with measurements. On the right is what the sample looks like on the alignment sheet.

Disks were placed onto a designed location on a grey masked cardstock, which was then placed in the tray of the DigiEye lighting box allowing for the sample to be at a 0° measuring angle from the SR. The A4 (210mm x 297mm) sized grey cardstock with CIELAB values of L*= 56.59, a*= 1.48, and b*= -2.24 was used to ensure each disk was aligned in the same location for measurement (Figure 71). Grey cardstock was chosen in order to match the base of the lighting cube. A 1 cm diameter circle was printed into the centre of the paper for each sample to be aligned inside of. The circle was aligned centrally on the A4 cardstock being 148 cm from the top and from the bottom of the page as well as 10 cm from the left and right side of the paper. While the DigiEye lighting box has a soft touch auto shut lock door reducing sample movement (VeriVide, 2008), prior to each measurement the view finder was used to ensure the aperture of the SR had approximate central alignment with the sample for increased precision.

5.2.3.2.2 45°/0 °Geometry

The spectroradiometer was configured on top of a copy stand with two lights to create an approximate 45°/0° geometry.

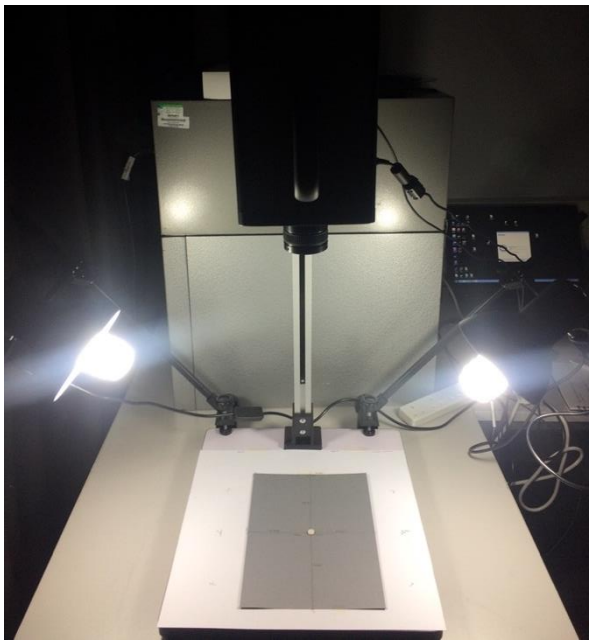

Figure 72: Konica Minolta Spectroradiometer with a copy stand set up to produce 45°/0° measurement geometry

The spectroradiometer was positioned at the 0° angle above the sample with the copy stand positioning light sources to provide 45° illumiation (Figure 72). The focal length was set at 0.5 and measurement angle was set to 1°. The distance between the objective lens and sample was 46.6 cm creating a measurement area of 7.25 mm or aperture diameter of 3.04 mm.

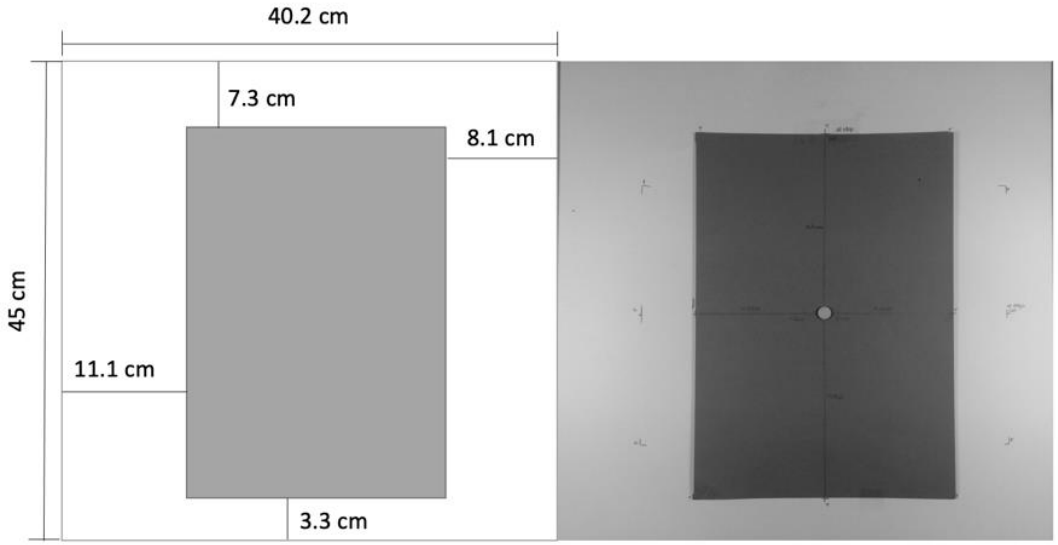

Figure 73: Grey alignment sheet arranged on a white alignment sheet to ensure sample was placed at a 0° angle from the spectroradiometer

A white cardstock with a length of 40.2cm and width of 45 cm was placed to cover most of the bottom of the copy stand. The samples were placed on the same grey cardstock alignment sheet as in Figure 71 in section 5.2.3.2.1. The grey alignment sheet was placed on top of a white alignment sheet 11.1 cm from the left, 8.1cm from the right, 7.3cm away from the top, and 3.3cm away from the bottom of the white cardstock alignment sheet in order to ensure the centre of each sample was placed at a 0° angle from the focal lens (Figure 73).

5.2.3.3 Vita EasyShade Advanced 4.0

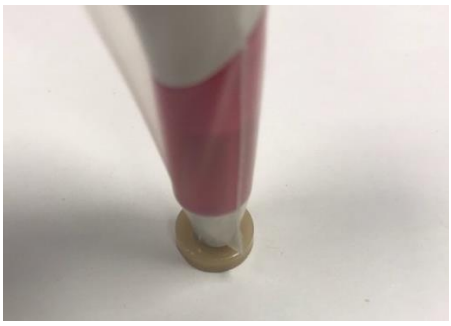

Figure 74: Vita EasyShade Advanced 4.0 taking measurement of one of the 52 custom disks

The probe of the EasyShade was placed approximately in the centre of each sample (Figure 74). Each measurement produces two readings: an inner reading and an outer reading based on which set of lights illuminates the sample.

5.2.3.4 DigiEye + Nikon D7000 (DN)

Figure 75: Custom disk on alignment sheet placed inside the DigiEye illumination cube

The samples were placed on the alignment sheet in Figure 71 in section 5.2.3.2.1 and placed inside the illumination cube (Figure 75). Alignment sheet allowed for 0° placement of the sample from the camera, producing a measurement geometry of diffuse/0°. The top of each sample was 49 cm away from the bottom of the camera lens. When the door to the illumination cube was shut, an image of the sample was captured. The pixel radius of the custom disks is 79.

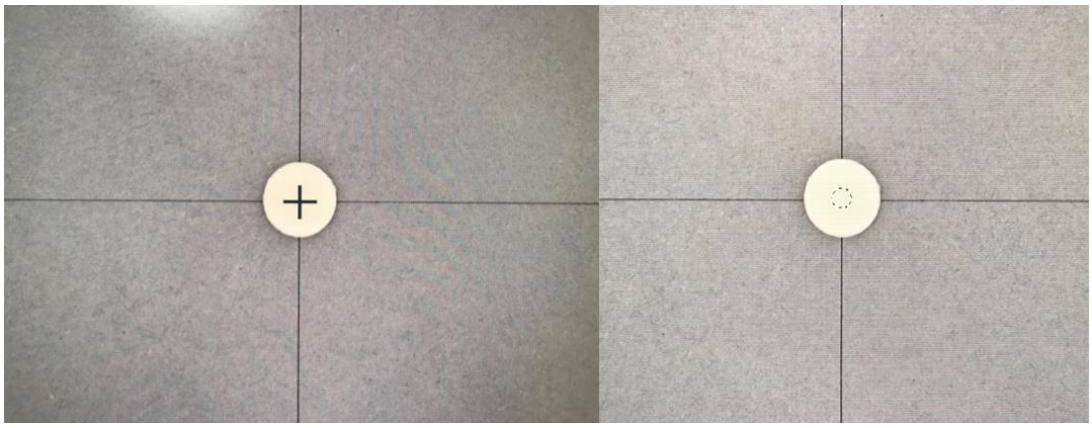

Figure 76: Image on the left is the DigiEye software measurement placement. Image on the right is the fixed pixel radius of 30 taking a colour measurement of that area only

To take a colour measurement, the cross is aligned in the centre of the sample (Figure 76). A fixed pixel circle radius of 30 was used to measure approximately the centre of each sample (Figure 76). This pixel radius provides a measurement size similar to the spectrophotometer measurement aperture of 3mm.

5.3 Data Analysis

All data was converted into CIEXYZ, CIELAB, and chromaticity values using methodology in section 2.6. The details of the method varied slightly depending upon the instrument. The Z-scores and different indices from each instrument were calculated using methods from section 2.4.1 (Z-scores) and section 2.8 (Indices). The correlation between indices and perceptual responses were compared using the statistical methods %WD and STRESS.

5.3.1 Scaling Analysis

For perceptual whiteness, the calculation method from section 2.4.1 was used where N is equal to 52 and the K is the mean rank order value of the responses from the 80 participants (UK and China) based on perceptual whiteness. For yellowness, calculation method from section 2.4.1 was used where N is equal to 52 and the K is the mean rank order value of the responses from the 40 participants (UK) based on perceptual yellowness.

The mean ranking for each tooth sample and the respective Z score for whiteness and yellowness are shown in Table 20. The sample with the highest Z score is considered perceptually to be whitest sample according to the participants (this is sample 0M1). The least white sample, with the lowest Z score, is 5M3. Based on perceptual yellowness, sample 5M3 is considered perceptually the most yellow with Z = 1.95; sample 0M1 is the least yellow with Z = -1.40. These Z scores are the reference standard to compare how accurate each device and each index is in evaluating tooth whiteness and tooth yellowness.

Table 20: Calculated Whiteness and Yellowness Z scores for the 52 samples based on the respective mean ranking value

Sample Name	Mean Whiteness Ranking	Whiteness Z score	Mean Yellowness Ranking	Yellowness Z score
0.5M1	3.76	1.61	45.42	-1.13
0M1	1.00	2.33	47.87	-1.40
0M2	2.05	2.04	47.08	-1.30
0M3	3.19	1.72	45.98	-1.19
1.5M2	16.29	0.53	35.02	-0.43
1M1	5.13	1.40	44.32	-1.03
1M1.5	10.28	0.91	38.24	-0.61
1M2	10.78	0.87	39.27	-0.68
2.5M2	23.38	0.15	27.06	-0.03
2L1.5	14.71	0.62	35.56	-0.46
2L2.5	21.49	0.25	27.37	-0.04
2M1	7.90	1.10	42.10	-0.86
2M2	18.83	0.39	31.36	-0.24
2M3	24.33	0.11	25.19	0.06
2R1.5	12.95	0.72	35.96	-0.48
2R2.5	12.56	0.75	37.35	-0.56
3.5M2	36.43	-0.51	18.46	0.41
3L1.5	21.64	0.24	28.62	-0.10
3L2.5	32.90	-0.32	18.51	0.40
3M1	18.38	0.41	33.38	-0.34
3M2	30.04	-0.17	22.22	0.21

3M3	33.50	-0.35	17.98	0.43
3R1.5	29.08	-0.13	23.55	0.15
3R2.5	37.08	-0.55	15.62	0.56
4.5M2	43.43	-0.96	10.45	0.90
4L1.5	40.36	-0.74	13.75	0.67
4L2.5	44.53	-1.05	8.66	1.04
4M1	36.44	-0.51	18.54	0.40
4M2	43.94	-1.00	9.94	0.93
4M3	46.56	-1.24	7.87	1.10
4R1.5	40.49	-0.75	12.86	0.73
4R2.5	45.71	-1.16	6.59	1.23
5M1	46.80	-1.27	8.97	1.01
5M2	47.83	-1.39	7.00	1.19
5M2.5	47.76	-1.38	6.65	1.22
5M3	51.78	-2.62	2.32	1.95
A1	8.61	1.04	39.37	-0.68
A2	23.78	0.13	28.06	-0.08
A3	28.14	-0.08	21.92	0.23
A3.5	35.21	-0.44	15.03	0.60
A4	42.88	-0.92	12.04	0.78
B1	6.25	1.26	43.35	-0.96
B2	16.28	0.53	33.07	-0.33
B3	32.31	-0.29	17.26	0.47
B4	40.51	-0.75	11.15	0.85
C1	12.76	0.74	38.52	-0.63
C2	26.18	0.02	26.50	0.00
C3	30.05	-0.18	21.69	0.24
C4	44.45	-1.04	12.55	0.75
D2	14.63	0.62	36.60	-0.52
D3	27.69	-0.06	24.09	0.12
D4	25.39	0.05	28.29	-0.09

5.3.2 Indices

From the colorimetric values, whiteness indices (WIC, WIO, WID) and yellowness indices (YIE313, YID1925, YIO) were calculated using methodology in section 2.8 for each sample from every device. An example of each of the indices calculated for each of the devices for one of the 52 samples, 0M1 is shown in Table 21. Appendix K-Appendix Q provides the calculated indices of the 52 samples for the each instrument.

Table 21: Indices calculated for sample 0M1 for each instrument

Instrument	WIC	WIO	WID	YIE313	YID1925	YIO	b*
SR	18.59	37.49	30.36	19.77	22.24	-36.88	7.57
SR45	17.05	35.17	29.55	19.53	22.10	-34.70	7.30
SPI	13.62	26.93	29.02	16.37	17.74	-26.51	4.79
SPE	16.73	30.59	29.01	17.20	19.25	-30.41	5.50
SP45	9.74	25.13	27.22	18.13	20.48	-24.89	5.90
ES	35.91	55.71	32.77	20.62	23.63	-55.22	8.99
DN	19.87	40.88	29.64	21.32	24.53	-40.34	8.91

5.3.3 % Wrong Decision (%WD)

The calculation followed the method in section 2.9.2. Percent wrong decision was calculated for each index in comparison to the whiteness Z score and then compared to the yellowness Z score.

Table 22: Visual perception Z scores for whiteness and yellowness for each sample compared to the calculated indices for measurement obtained with the SR

Sample Name	Whiteness Z Score	Yellowness Z score	WIC	WIO	WID	YIE 313	YID 1925	YIO	b*
0.5M1	1.61	-1.13	-8.32	21.70	23.69	26.21	30.71	-20.65	11.66
0M1	2.33	-1.40	18.59	37.49	30.36	19.77	22.24	-36.88	7.57
0M2	2.04	-1.30	9.83	32.28	27.47	22.15	25.62	-31.67	9.13
0M3	1.72	-1.19	-5.75	23.01	24.21	25.55	29.89	-22.04	11.22
1.5M2	0.53	-0.43	-50.97	-5.21	12.68	35.34	43.19	6.57	17.15
1M1	1.40	-1.03	-20.13	14.42	20.47	28.92	34.43	-13.29	13.37
1M1.5	0.91	-0.61	-39.97	2.83	15.24	33.64	40.84	-1.51	16.47
1M2	0.87	-0.68	-41.42	2.56	15.90	33.83	40.69	-0.95	16.54
2.5M2	0.15	-0.03	-56.87	-11.01	9.42	36.37	45.42	11.82	17.60
2L1.5	0.62	-0.46	-44.69	-0.70	14.85	34.04	41.16	2.19	16.41
2L2.5	0.25	-0.04	-71.12	-14.74	8.90	40.35	49.37	16.77	20.76
2M1	1.10	-0.86	-29.61	6.35	17.81	30.01	36.14	-5.34	13.61
2M2	0.39	-0.24	-63.54	-11.07	9.68	38.78	47.68	12.71	19.76
2M3	0.11	0.06	-66.71	-15.19	7.44	39.10	48.83	16.33	19.69
2R1.5	0.72	-0.48	-37.26	2.18	14.61	32.50	39.92	-1.36	15.47
2R2.5	0.75	-0.56	-43.97	-1.53	13.59	33.88	41.55	2.59	16.29
3.5M2	-0.51	0.41	-67.23	-21.56	5.06	37.55	48.12	21.63	17.52
3L1.5	0.24	-0.10	-43.24	-3.21	14.24	32.51	39.67	4.22	14.82
3L2.5	-0.32	0.40	-76.71	-21.85	6.46	40.41	50.19	23.34	19.90
3M1	0.41	-0.34	-47.44	-6.69	12.63	33.26	40.96	7.53	15.14
3M2	-0.17	0.21	-68.60	-18.56	6.29	38.85	48.94	19.35	19.05
3M3	-0.35	0.43	-72.45	-20.79	5.41	39.70	50.08	21.63	19.58
3R1.5	-0.13	0.15	-53.53	-13.13	7.63	34.86	44.66	12.96	16.16
3R2.5	-0.55	0.56	-74.98	-25.35	1.90	40.16	52.04	25.24	19.72
4.5M2	-0.96	0.90	-73.41	-27.95	2.79	38.14	49.49	27.62	17.23
4L1.5	-0.74	0.67	-79.89	-30.50	-0.37	40.81	53.57	29.95	19.73
4L2.5	-1.05	1.04	-70.45	-24.71	4.75	37.55	48.06	24.81	16.99
4M1	-0.51	0.40	-46.93	-13.02	8.57	31.81	41.07	12.35	13.52
4M2	-1.00	0.93	-82.79	-32.77	0.29	40.70	53.04	32.49	19.07
4M3	-1.24	1.10	-83.87	-33.21	-1.25	41.54	54.60	32.65	20.05
4R1.5	-0.75	0.73	-59.80	-22.16	3.48	35.07	46.46	20.94	15.47
4R2.5	-1.16	1.23	-79.67	-33.40	-1.39	40.00	53.23	32.27	18.59
5M1	-1.27	1.01	-56.22	-19.55	7.57	32.92	42.18	19.20	13.51
5M2	-1.39	1.19	-80.81	-34.66	-0.79	39.66	52.55	33.69	17.92
5M2.5	-1.38	1.22	-82.39	-35.39	-1.58	40.29	53.53	34.34	18.49

5M3	-2.62	1.95	-128.97	-62.21	-12.98	51.07	68.56	61.24	25.64
A1	1.04	-0.68	-39.21	4.45	16.41	33.62	40.38	-2.83	16.55
A2	0.13	-0.08	-60.06	-13.92	6.96	37.32	47.38	14.28	18.31
A3	-0.08	0.23	-68.45	-18.22	5.67	39.18	49.59	18.89	19.50
A3.5	-0.44	0.60	-78.10	-26.29	2.47	40.70	52.23	26.60	19.98
A4	-0.92	0.78	-73.90	-27.05	2.99	38.62	49.91	26.91	17.81
B1	1.26	-0.96	-29.44	9.64	20.63	30.38	35.44	-7.77	14.04
B2	0.53	-0.33	-57.09	-8.06	11.91	36.76	44.79	9.71	18.09
B3	-0.29	0.47	-83.39	-26.02	3.63	42.26	53.16	27.26	21.36
B4	-0.75	0.85	-98.01	-35.81	-1.21	45.48	58.14	36.73	23.45
C1	0.74	-0.63	-42.06	0.68	17.74	32.39	38.12	1.32	14.87
C2	0.02	0.00	-56.08	-11.59	10.61	35.30	43.70	12.55	16.41
C3	-0.18	0.24	-65.93	-17.56	9.47	36.92	45.45	18.90	17.00
C4	-1.04	0.75	-74.41	-25.77	6.04	38.01	47.75	26.57	16.99
D2	0.62	-0.52	-43.81	-2.36	15.77	32.49	38.94	3.86	14.75
D3	-0.06	0.12	-60.44	-16.34	7.88	36.02	45.48	16.75	16.64
D4	0.05	-0.09	-66.12	-14.22	10.56	38.03	46.29	16.10	18.44

Table 22 presents the calculated index values for each sample from measurements collected from the SR compared to whiteness and yellowness Z-scores. A total of 1326 comparisons can be made per index compared to one set of z-scores.

Table 23: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the SR compared to whiteness and yellowness Z scores

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1146	1326	13.57	161	1326	12.14
WIO	1228	1326	7.39	77	1326	5.81
WID	1203	1326	9.28	95	1326	7.16
YIE313	275	1326	20.74	1079	1326	18.63
YID1925	210	1326	15.8	1146	1326	13.57
YIO	108	1326	8.14	1235	1326	6.86
b*	375	1326	28.28	976	1326	26.40

Table 23 displays the number of correct comparisons each indices calculated from measurements collected from a SR made compared to the perceptual whiteness as well as perceptual yellowness of the 52 samples.

As mentioned in section 2.9.2, when a yellowness index is compared to perceptual whiteness it is actually the inverse that is displayed in the %WD location. This is the same when a whiteness index is compared to perceptual yellowness. This was repeated for each indices from each instrument used in the study. Appendix R contains the number of correct comparisons were made by each instrument and specific index compared to each of the visual perception reference standards.

5.3.4 STRESS Analysis

STRESS between perceptual responses (Z scores) and index was calculated using equations in section 2.9.3. This was repeated for each index calculated by each instrument for both perceptual responses.

5.4 Results

Table 24 and

Table 25 show the %WD and STRESS values for each instrument and indices compared with the visual whiteness data.

Table 24: % Wrong Decision of instruments and computed indices compared to the visual perception (z score) of 80 total Participants from China and the UK for the evaluation of tooth Whiteness

Instrument	WIC	WIO	WID	YIE313	YID1925	YIO	b*
SR	13.6	7.4	9.3	20.7	15.8	8.1	28.3
SR45	12.8	7.4	8.3	18.6	14.4	7.6	25.7
SPI	15.5	8.7	9.3	21.3	16.9	10.0	26.9
SPE	14.4	8.4	8.3	18.6	15.1	9.4	24.7
SP45	13.4	7.5	17.3	13.7	13.7	8.5	22.5
ES	13.9	12.1	11.4	16.9	14.0	12.1	21.4
DN	14.4	7.2	10.3	22.0	16.7	8.4	30.9
Mean	**14.0**	**8.4**	**10.6**	**18.8**	**15.2**	**9.2**	**25.8**

Table 25: STRESS analysis of instruments and Indexes compared to the Z scores of 80 total Participants from China and the UK for the evaluation of tooth Whiteness

Instrument	WIC	WIO	WID	YIE313	YID1925	YIO	b*
SR	16.4	10.2	11.3	21.4	17.8	11.4	27.8
SR45	15.8	9.8	10.3	20.1	16.4	11.0	25.9
SPI	17.2	11.3	11.2	21.9	18.5	12.5	27.1
SPE	16.3	11.0	10.7	19.9	16.7	12.0	25.3
SP45	15.1	9.5	8.8	18.4	15.2	10.6	23.2
ES	15.7	14.3	13.8	18.2	16.2	14.6	22.4
DN	16.5	10.1	11.7	22.1	18.0	11.3	29.4
Mean	**16.1**	**10.9**	**11.1**	**20.3**	**17.0**	**11.9**	**25.9**

If we consider the %WD data in Table 24 and calculate the mean performance of all the instruments, then the best preforming index for predicting changes in perceptual whiteness is WIO (mean=8.4). Although YIO (mean=9.2) and WID (mean=10.6) also perform well. Indices WIC (mean=14.0), YIE313 (mean=18.8), YID1925 (mean=15.2) and b* (mean=25.8) have higher values indicating these equations are not suitable for predicting whiteness in comparison to WIO, WID and YIO. The poor performance of these indices confirms that index needs to be used on the type of material for white it was intended to be valid (Joiner and Luo, 2017; Mohan *et al.,* 2008). WIO, WID, and YIO were developed for the use in dentistry and this could be the reason these indices outperform the other indices (Sullivan *et al.*, 2019; Luo *et al.*, 2009; Pérez *et al.*, 2016). This is further confirmed by the mean STRESS results in

Table 25. WIO (mean=10.9) and WID (mean=11.1) are the best indices for predicting whiteness. Although, YIO (mean=11.9) also performs well.

Table 26: The squared STRESS ratios of the mean index values comparing the indices for predicting whiteness

	WIC	WIO	WID	YIE313	YID1925	YIO	b*
WIC		2.21	2.11	0.63	0.90	1.83	0.39
WIO			0.96	0.29	0.41	0.83	0.18
WID				0.30	0.43	0.87	0.18
YIE313					1.43	2.89	0.62
YID1925						2.02	0.43
YIO							0.21
b*							

The squared STRESS values of the mean index value of all the instruments compared to other mean index values are shown in Table 26. The F_c for this experiment is (0.95,51,51) producing a critical F value of 1.60 producing a confidence interval between 0.625 and 1.60 (Purdue, 2015). Any ratio outside this interval, shown in red in Table 26, indicates that the two indices are statistically different. WIO is statistically different from all other indices except for YIO (0.83) and WID (0.96). WID and YIO are also not statistically different (0.87).

Table 27 and Table 28 show the %WD and STRESS values respectively for each instrument and indices compared with perceptual yellowness data. This table includes the calculated mean performance of all the instruments.

Table 27: % Wrong Decision of instruments and Indexes compared to the Z scores of 40 UK participants for the evaluation of tooth yellowness

Instrument	WIC	WIO	WID	YIE313	YID1925	YIO	b*
SR	12.1	5.8	7.2	18.6	13.6	6.9	26.4
SR45	11.7	5.3	6.3	17.0	12.4	6.6	23.8
SPI	13.7	6.9	7.3	19.5	15.1	8.5	25.1
SPE	12.7	6.8	6.3	17.0	13.3	8.1	23.1
SP45	12.1	5.9	4.8	15.9	12.1	7.1	21.1
ES	14.0	12.4	11.2	16.3	13.6	12.4	20.7
DN	12.7	6.3	8.1	20.1	14.5	7.5	28.9
Mean	**12.7**	**7.1**	**7.3**	**17.8**	**13.5**	**8.2**	**24.2**

Table 28: STRESS analysis of Devices and Indices compared to the mean rank order of visual perception of yellowness from 40 UK participants

Instrument	WIC	WIO	WID	YIE313	YID1925	YIO	b*
SR	11.6	8.2	10.3	15.8	13.3	8.3	21.8
SR45	11.1	8.0	10.0	14.7	12.2	8.1	20.1
SPI	12.7	8.6	11.0	17.3	15.0	8.9	22.4
SPE	12.2	8.9	10.9	15.4	13.6	9.0	20.8
SP45	12.0	9.2	11.2	14.8	13.1	9.2	19.5
ES	13.3	13.6	15.0	15.3	15.0	13.3	19.3
DN	11.8	7.7	10.5	16.7	13.8	8.1	23.7
Mean	**12.1**	**9.2**	**11.3**	**15.7**	**13.7**	**9.3**	**21.1**

Similar to the perceptual whiteness results for %WD, the best performing indices is WIO (mean=7.1). WID (mean=7.3) and YIO (mean=8.2) performed just as well in predicting changes in perceptual yellowness. In terms of STRESS, all three of these metrics; WIO (mean=9.2), YIO (mean=9.3) and WID (mean=11.3) perform well.

WIO, YIO and WID produce overall lower %WD and STRESS values for evaluating yellowness compared to whiteness. It is unexpected that this is the case for WIO and WID, as these indices were developed for the evaluation of whiteness.

Table 29: The squared STRESS ratios of the mean index values comparing the indices for predicting yellowness

	WIC	WIO	WID	YIE313	YID1925	YIO	b*
WIC		1.74	1.16	0.59	0.78	1.70	0.33
WIO			0.66	0.34	0.45	0.98	0.19
WID				0.51	0.67	1.47	0.29
YIE313					1.31	2.87	0.56
YID1925						2.19	0.42
YIO							0.19
b*							

Similar to perceptual whiteness, WIO, WID and YIO are statistically indistinguishable with squared STRESS ratios of WIO vs YIO (0.98), WIO vs. WID (0.66) and YIO vs. WID (1.47) shown in Table 29. It is interesting

that WID and WIC are not statistically different. The index WIC, however, has a higher STRESS value meaning it has worse correlation with perceptual yellowness. The performance of STRESS and %WD for the YIO index further validates of the development of index from chapter 3 for evaluation of perceptual tooth yellowness (Sullivan *et al.*, 2019).

While indices WIC, YIE313, and YID1925 have been used for evaluating changes in tooth colour (Luo *et al.,* 2009; del Mar Pérez *et al.,* 2016), evidence is quite strong that none of the indices correlate well with either perceptual whiteness or yellowness. This is not surprising as they were not developed for use in dentistry but for other materials (i.e. textiles, paints, oils and plastics) (ASTM E313-05, 2005; Hunter, 1981). As mentioned in chapter 3, this study further confirms with %WD results (b* mean=24.2) and STRESS (mean=21.1) that this metric should not be used to evaluate yellowness. These results confirm that changes in tooth colour is more complex than just shifting the blue-yellow scale.

While WIO, YIO and WID are statistically indistinguishable, there is inconsistencies between on what is considered the ‘best performing’ instrument depending on the index used. According to WIO, DN instrument gives the best performance according to %WD (Table 24) and the SP45 instrument gives the best performance according to STRESS (

Table 25) for whiteness. For yellowness, the SR45 instrument gives the best performance according to %WD (Table 27) and the DN instrument gives the best performance according to STRESS (Table 28). Like WIO, when using YIO and WID, according to STRESS the best instrument is an SP45 for evaluating whiteness (Table 10). However, according to YIO and WID %WD (Table 9) the SR45 instrument gave the best performance. According to WID %WD, SPE performed equally as well as SR45. For yellowness, when using YIO according to %WD the SR45 is the best instrument to use while STRESS results are similar to WIO results which suggests the best instrument is DN. Using WID for evaluating yellowness, %WD results suggest using an SP45 while STRESS results indicate DNA is the best method.

All three metrics performed better for evaluating yellowness than whiteness. Based on yellowness data, the 'best performing' device is the DN as it produced the lowest overall STRESS value for both WIO and YIO and therefore is used to compare all other instruments too. It produce the lowest %WD for WIO for whiteness as well and is therefore being used to compare all other instruments to.

Table 30: Squared STRESS ratios for each instrument compared with the instrument (DN) based on the indices WIO, WID, and YIO

	WIO		WID		YIO	
	Whiteness	Yellowness	Whiteness	Yellowness	Whiteness	Yellowness
SR	1.02	1.12	0.93	0.97	1.00	1.07
SR45	0.95	1.07	0.77	0.91	0.95	1.01
SPI	1.27	1.25	0.90	1.09	1.21	1.22
SPE	1.18	1.32	0.83	1.08	1.13	1.25
SP45	0.88	1.42	0.56	1.13	0.88	1.31
ES	2.00	3.08	1.37	2.04	1.67	2.74
DN	1.00	1.00	1.00	1.00	1.00	1.00

Table 30 shows the squares of the STRESS ratios for each of the instruments where the denominator of the ratio is the STRESS for DN. Values in red indicate outside of the confidence interval of 0.625 to 1.60 and indicate statistical difference. DN is statistically indistinguishable from

all other instruments except for the EasyShade. This could be due to the perceptual data versus the instrument used as it is not statistically different for any other index. WID displays different results with SP45 being the only instrument to be statistically different than DN. If the SP45 was selected as the as the 'best' performing instrument for other devices to compared too, the results differ. SP45 was chosen as the instrument produced the lowest STRESS value for whiteness for all three metrics. This is the only instrument all three metrics agreed upon.

Table 31: Squared STRESS ratios for each instrument compared with the instrument (SP45) based on the indices WIO, WID, and YIO

	WIO		WID		YIO	
	Whiteness	Yellowness	Whiteness	Yellowness	Whiteness	Yellowness
SR	1.16	0.79	1.66	0.86	1.14	0.82
SR45	1.07	0.75	1.38	0.80	1.07	0.77
SPI	1.44	0.88	1.62	0.96	1.37	0.93
SPE	1.34	0.93	1.48	0.95	1.28	0.95
SP45	1.00	1.00	1.00	1.00	1.00	1.00
ES	2.27	2.17	2.46	1.80	1.89	2.09
DN	1.13	0.70	1.79	0.88	1.13	0.76

Table 31 shows the squares of the STRESS ratios for each of the instruments where the denominator of the ratio is the STRESS for SP45. Values in red indicate statistical difference, values falling outside of the confidence interval of 0.625 to 1.60. Results indicate that ES performs statistically worse than SP45 for all three indices for both whiteness and yellowness. It is interesting to note that when using WID, there are more instruments which are statistically different than SP45 for evaluating whiteness compared to indices WIO and YIO. This indicates that the index is not as robust as WIO and YIO.

5.5 Discussion

In general, indices WIO, WID and YIO performance for evaluating yellowness and whiteness was strong producing the lowest overall values for STRESS and %WD. The performance of these indices is not surprising

since all three metrics were optimised from colour measurements of Vita shade guides and perceptual evaluations of tooth whiteness and yellowness (Luo *et al.*, 2009; Sullivan *et al.*, 2019; Pérez *et al.*, 2016). This study evaluated these metrics using physical measurements and perceptual data about custom disks created by Vita that are based on shades from their shade guides. WIO, WID and YIO were found to be statically indistinguishable for whiteness and yellowness assessment.

WIO, WID and YIO were developed to be optimised using VITA shade guides. The samples used in this study were made by VITA and were based on shade from the shade guide tabs. While these indices perform well, results could differ if shade guides made by different companies were used such as Chromascope, Filtek Z250, Premise, Charisma, Grandio, or made of different materials such as hybrid shade guides and nano hybrid shade guides (Cal *et al.* 2004; Yamanel *et al.* 2010).

All the STRESS values for WIO, YIO and WID both whiteness and yellowness are all in the range of 7.7-15. According to Kruskal (1964), a STRESS value of 5 is considered good correlation, 10 is fair and 20+ is poor. These values are in the area of having fair correlation with perceptual data. It could be argued these STRESS are actually low and produce good correlation for colour difference indices. A study conducted by Melgosa *et al.* (2008), comparing three different colour difference formulas (CIE94, CIELAB, and CIEDE2000) all the results produced STRESS values of 18.85 and higher. A value of 18.85 was considered to produce good results. This suggests that for a colour difference evaluating index such as WIO and YIO, a value of 7 is really good correlation to the data. Ideally, the best method would produce a STRESS value of 0. This may not be possible due to the variability of indices, differences in perceptual results, and edge-loss. It is known that visual perception is imperfect and heavily subjective influencing perceptual results (Hammad, 2003).

It is unexpected that whiteness indices WIO and WID produced better correlation with perceptual yellowness than with perceptual whiteness. These results could be a reflection on perceptual data versus the indices used; suggesting it may be easier to visually assess yellowness attributes of teeth over whiteness.

The agreement with visual assessment was the basis for accuracy of instruments when measuring tooth coloured samples. Results suggest that there is no 'best' performing instrument, but found that all other instruments and measurement geometries were not statistically different than DN except for the ES (Table 29). These results could be due to the fact that the differences in measurement values between the devices is so small that it does not affect the index value enough to produce statistical differences. It is not surprising that there is no statistical difference between SP devices and SR devices as white data from a spectrophotometer is used to convert irradiance data into reflectance data. It is interesting that the digital camera is not statistically different than SP. However, this could be due to the fact that the digital camera software uses spectrophotometer measurements of a colour chart in order to calibrate the camera and convert RGB data into XYZ data.

Table 32: Description of instrument characteristics and corresponding WIO STRESS and %WD results for perceptual tooth yellowness

Instrument	Illumination Angle/ Measurement Angle	Illumination Area	Measurement Area	Contact vs. Non-Contact	STRESS (WIO)	%WD (WIO)
SR	Diffuse/0°	≈ all	2.83mm	Non	8.20	5.81
SR45	45°/0°	≈ all	3.04 mm	Non	8.00	5.28
SP (SCI)	Diffuse/8'°	6mm	3mm	Contact	8.65	6.94
SP (SCE)	Diffuse/0°	6mm	3mm	Contact	8.88	6.79
SP45	0°/45°	6mm	4mm	Contact	9.23	5.88
ES	Unknown	Unknown	1mm	Contact	13.58	12.37
DN	Diffuse/0°	≈ all	≈ 3mm	Non	7.74	6.33

Table 32 describes each instruments measurement characteristics compared to the STRESS and %WD for WIO for yellowness. There is no consistency with the results. In general, measurements in which illuminated the entire sample produced better STRESS results. The 45/0 measurement geometry produced better %WD results in comparison to samples with diffuse illumination. According to Paul *et al.* (2002), the 45° illumination/0° detection is the most easily replicated practically in a dental office. ES was the only device that was considered statistically different and produced the highest STRESS and %WD. This could be due to the fact that it has an unknown illumination area/measurement area. It was found to be statistically different than the DN. This could be due to the fact that the instrument does not actually collect a colour measurements. It is a device that compares the collected measurement to the closest shade guide tab from a database and produces a set of colorimetric values and spectral data that corresponds to that specific shade guide. The issue with this is that there is an assumption that the database values are 'correct'. As well as it could be selecting an inaccurate shade from the database.

While the results indicate that spectrophotometers are not statistically different than spectroradiometer or digital cameras, they are contact based devices. For clinical dental applications*, in vivo* measurements are required. This presents particular difficulties for many spectrophotometers with the exception of dental specific spectrophotometers explicitly designed for in vivo measurements such as the ES used in this study (Rauf, 2020). However, ES produced the worst agreement with visual perception compared to the other instruments. In general, non-contact devices produced lower STRESS values and %WD results than the contact based devices with the exception of %WD for SP45 (Table 32). However, in practice the DigiEye lighting cube used for the DN set up could not be used for *in vivo* measurements. The requirement for external lighting might introduce additional error when an imaging system is used in clinical setting. Besides the fact that the digital cameras produced the

lowest STRESS values for yellowness, there are benefits to using a digital camera in a clinical setting versus a spectroradiometer. Digital cameras are readily available, cheap, and easy to use compared to spectroradiometers (Guan *et al.,* 2005; Pan and Westland, 2018; Lee *et al.,* 2011; Wee *et al.,* 2006). In addition they provide visual evidence of efficacy of aesthetic treatments and allows for the production of a database of images archiving an individual's health (Schmidseder, 2000; Brook *et al.*, 2007; Pan and Westland, 2018; Mohan *et al.,* 2008).

The investigation of devices for research and development (R&D) purposes is equally as important as the need for devices to be used in a clinical setting (i.e. *in vivo* measurements). In R&D, instruments will be used for investigating the efficacy of toothpastes and whitening treatments, development of new products, and generating claims based on these products such as 'reduces up to 15 years of discoloration' or '4 shades visibly whiter'. *In vivo* measurements are not required for the development of a new product until the clinical trial stage. Therefore, devices such as the DigiEye and spectrophotometers could easily be used for R&D purposes. Aside from the EasyShade, any device could be used for evaluating the efficacy of new toothpaste products or whitening products in research with the use of WIO or YIO to evaluate changes in tooth colour.

Academic research from Bahannan (2014), Da Silva *et al.* (2008), and Paul *et al.* (2002) suggest that visual assessment of tooth colour is regarded as inaccurate compared to instrumental devices even with trained technicians and therefore should not be used as the gold standard for comparing devices too. These studies use either a small number of technicians (i.e. three or less) or inexperienced dental students which is not representative of the number of technicians that do visual assessments daily. Universally if visual assessment of tooth colour produced unsuitable results clinically, then it would not be used. Not only is the clinician visual assessment still the final arbitrate for dental

restorations and cosmetic procedures, but the patient must also deem results satisfactory or not by visually evaluating their own dentine results. Additionally, how others perceive an individuals' smile is via visual perception (Montero *et al.*, 2014; Kershaw *et al.*, 2008). Therefore, visual assessment is used as the 'gold standard' for which instrumental measurements should be compared to.

In the light of the challenges identified, a novel method has been introduced whereby accuracy is expressed in terms of the ability of the instrument to give measurements that 'agree with' visual assessments. Other approaches may be possible. Paravina *et al*. (2021) conducted a study which investigated a custom-tailored calibration method for the needs in dentistry instead of calibrating with an opaque ceramic plaque. They found that when implementing correction factors to the calibration process there was a 70% reduction in colour difference values (∆E00 and ∆Eab) for tooth measurements taken with a spectrophotometer (d/8°) and a spectroradiometer (45°/0°). This approach could be applied to digital cameras by developing a tooth colour chart, a method which has been applied to tongue images in order to compare images captured in different environments (Zhang *et al.*, 2018). If calibration corrections were used for every instrument, the performance of each device and index could vary.

This experiment only covers a small selection of colour measurement devices based on availability with their standard calibration processes. Other devices that were not used in this study specifically designed for dentistry include spectrophotometers crossed with digital cameras such as CrystalEye, the 3Shape TRIOs device and Sopro 717. Other devices that should be tested in the future should include a range of cell phones for the evaluation of at home whitening products. However, cell phones produce another set of challenges such as process of capturing image, differing RGB filters and JPG conversion errors, type of phone, lighting, and angle of which image was captured (De Bragança *et al.,* 2021). These

newer devices could produce different results and might be more accurate than the selection of instruments in this study.

5.6 Conclusions

This study raised the question of how to assess accuracy when using instruments to measure tooth colour. A novel approach to the assessment of accuracy is introduced whereby accuracy is expressed by the instrument's ability to 'agree' with visual perception. WIO, WID and YIO were found to be statistically indistinguishable for evaluating tooth whiteness and yellowness. However, all three indices performed better for evaluating tooth yellowness versus tooth whiteness which raises questions in regards to what is the 'best' method for visual assessment of tooth colour.

There was no one 'best performing' device. While the digital camera was used for comparison, most devices were found to not be statistically different than the DN besides the ES. ES was found to be statistically different and therefore may not be suitable for evaluating perceptual changes in whiteness or yellowness. This indicates that the device might not be as important as the index used. SP45 was also used for comparison and found that WID is not as robust of an index as WIO or YIO when varying instruments.

6 Overall Discussion

Accuracy of tooth colour measurement was explored in this thesis; investigating the relationship between visual perception and instrumental measurement of tooth colour. Experiments in this thesis have highlighted issues in different areas of this relationship including the correlation issues within the tooth colour gamut, multiple indices are capable of being developed due to the lack of degrees of freedom between tooth colour values, edge-loss, and there is no ‘gold’ standard method or ‘best performing’ device in which data can be compared to. This includes measurement geometry, instrument, contact vs. non-contact, visual assessment and index used. The goal of colour measurement devices is to relate numerical values to which visual perception. The best that can be done for defining accuracy is evaluating the correlation between instruments, indices and visual perception. Despite highlighting the many problems with tooth colour measurement, a yellowness index was developed for evaluating perceptual yellowness in teeth for use in dentistry and a new method for evaluating the accuracy of indices and instruments is based on the correlation to visual perception.

A new yellowness index, YIO was optimised for evaluating changes in perceptual yellowness of teeth. Metric b* was found to be unsuitable for evaluating changes in tooth yellowness. In comparison to all other indices WIO, WID and YIO were found to be statistically indistinguishable for evaluating tooth whiteness and yellowness. WIO and YIO were found to be more robust than WID when comparing instruments to visual perception. One issues is that when comparing WIO and YIO to uncorrelated data, the indices performed poorly. The other indices were not evaluated against uncorrelated data and potentially could perform better than WIO and YIO. This raises the question if a more robust index could be developed.

Principal component analysis should be investigated for use of developing a more robust index for tooth colour. Principal component analysis is a statistical method for understanding correlated relationships by creating artificial uncorrelated variables optimized to maximise the variation in a set of data for each component (Hess and Hess, 2018; Abdi and Williams, 2010; Westland et al., 2012). This method might be able to isolate the different attributes of tooth colour that are not accounted for in the indices currently developed for tooth colour analysis.

WIO, WID, and YIO performed better for evaluating tooth yellowness versus tooth whiteness which raises questions in regards to what is the 'best' method for visual assessment of tooth colour. Indices currently developed for tooth colour analysis are based on visual perception of either 'whiteness' or 'yellowness'. Yellowness is an easier attribute to assess as it is universally more understood than 'whiteness'. Whiteness is a difficult to assess considering teeth cannot be a 'perfect' white. Whiteness is multi-dimensional attribute to assess as the hue considered closest to 'white' is assessed prior to assessing lightness. This could be the reason WIO performs slightly better than YIO as it takes into consideration other hues besides yellow. However, both indices are solely focused on these two attributes of tooth colour but red hues, blue hues, and green hues can influence tooth colour. Red hues specifically with heavily stained teeth. Future work should investigate visually evaluating teeth based on perceptual health or attractiveness for example, in order to include and or prioritise other tooth attributes which might not be considered as important when evaluating based on 'whiteness' and 'yellowness' attributes alone.

Another issue that is mentioned in dental research is measurement devices are either subjected or not to edge-loss. A spectroradiometer, spectrophotometer, digital camera, a dental spectrophotometer (Vita EasyShade) were compared to evaluate the amount of edge-loss caused

by instrument, sample size, translucency and finish. All of these instruments produced different CIELAB values for the same sample. While this could be due to edge-loss, the differences could also be due to the differences in measurement geometry, illumination strength, or if the device has contact with the sample or not. The SR, SP and ES were more affected by the degree of translucency than the DN. The DN device is more influenced by the finish of the sample. The size of the sample affected all of the devices differently. Results determined that both contact-based devices and non-contact based devices are effected by edge-loss. The a* and b* values are equally as influenced as L* by differing translucency, size of sample and device used. It is unclear if any of the devices produce more edge-loss compared to another device without knowing the 'true' value of the sample. The ES produced the highest L* values in comparison to the SR, SP, and DN. In comparison to visual perception the ES was found to be statistically different to every other device and deemed unsuitable for evaluating changes in tooth colour. This indicates that in the edge-loss study that the ES over-estimated the colour values of the sample in comparison to the other devices. In the edge-loss study, the averaged ΔE00 value between SR, SP and DN instruments is 4.2 and an averaged ΔEab*=5.6. These values are outside the perceptibility threshold (PT) and acceptability threshold (AT) of PT Δ E00=0.8, PT Δ Eab*=2, AT Δ E00=3.7 and AT Δ Eab*=1.2 meaning that these differences are noticeable (Khashayar *et al.*, 2014; Paravina *et al.*, 2015). When evaluating their accuracy against visual perception, these devices were found to be statistically undisguisable. While edge-loss does affect the measurement, it might not be as important if the same instrument is used for every measurement. Results from the accuracy chapter indicate that the instrument might not be as important as the index used.

Indices YIO and WIO have outperformed all other indices relating instrumental measurements to visual perception. While these indices can evaluate perceptual changes in tooth colour for use such as evaluating the efficacy of tooth whitening produces, it is unknown the amount of difference between index values is perceptibility. Currently studies have only investigated acceptability and perceptibility thresholds for colour difference (∆Eab* and ∆E00) values, but not for index values. Development of perceptibility and acceptability thresholds are essential for the assessment of efficacy of tooth whitening product development, as well as at-home and in-office procedures (i.e. whitening, restorations).

One issue is that custom made flat disks were used for experiment in chapter 5. While flat porcelain disks have been used for dental research as they provide a flat measurement surface and reduce errors (i.e. stray light) in colour measurement, the disks do not provide an accurate representation of tooth structure. (Seghi, 1990; Paravina *et al.,* 2021; Pérez *et al.,* 2016). Teeth are curved, vary in size (i.e. incisors, canines, molars), glossy, and translucency. If the samples were more like shape of a tooth, results would differ as curved surfaces would introduce physical measurement errors such as stray light for the contact-based devices. Even with the flat surfaces, the samples have two translucent layers similar to teeth. These characteristics of tooth colour measurement will effect instrumental measurements and may affect the way teeth are visually seen.

In addition, these disks were developed to match standardised Vita shade guide tabs. While shade guides provides samples that correspond to a range of human tooth colour in order to increase chances for shade matching, there is a lack of standardization between tabs (Chu et al., 2010). Clinically, shade guides are still the oldest and most frequently used method for tooth colour assessment and communication in dentistry (Paravina, 2008; Mohan et al., 2008; Koumpia et al., 2018). While there a

few different companies that made shade guides and they are imperfect in uniformity, shade guides are the only ‘standard’ that are available for use in research that mimic tooth colour gamut (Paul et al., 2002). The samples used in this thesis are based on a combination of different shade guides Vita bleachedguide 3D Master, Vita Classical, and Vita Toothguide, Vita extended Bleachedguide 3D Master to increase the range of tooth colour covered. However the shade guides still do not cover the entire range of human tooth colour (Klotz et al., 2018). Shade tabs are not representative of all teeth and do not account for dental discoloration from various diseases or cavities. These samples also do not take into account the mixture of natural teeth and dental restorations. Shade tabs also have different optical properties than natural teeth, which could influence the performance of the indices and devices on measuring natural teeth (Klotz et al., 2018). All of these aspects could affect the accuracy of the index and measurement instrument in clinical settings. Further research should be conducted on ways to expanding standardised samples.

Despite the deficiencies of shade guide tabs, they were used in the research conducted in this thesis. Shade guide tabs are a pragmatic standard chosen due to the fact that there are no other dental standards available. In addition, shade guide tabs are still extensively used in clinical dentistry. If shade guide tabs were statistically producing unsatisfactory results, they would not be used in practice. Multiple different brands of shade guide tabs exist, but VITA was chosen as their shade guide tabs have been used in industry for the last 50 years and is widely used well-known brand (VITA North America, 2022).

The need for these devices to be used in a clinical setting is important. A standard spectrophotometer is not able to measure in-vivo. While not useful for in-vivo, these devices could still be suitable for research and development teams. In two of the experiments in this thesis the Digi-eye was used which is an enclosed system with a digital camera providing

consistent, even, standardised illumination. In clinical practice, this device would not be able to be used suggesting that it might not be suitable for the evaluation of teeth. This means that further research needs to be done on the use and accuracy digital cameras such as the TRIOS device. As technology advances in colour measurement, the accuracy of new devices and indices will need to be evaluated against visual perception.

Results from this thesis either has been published previously or aims to be published in the future. The yellowness index has already been published in the Journal of Dentistry and has been used in clinical studies conducted by Colgate in which the efficacy of tooth whitening products was investigated (Sullivan, 2019). The Colgate advertisement “reverses up to 15 years of discoloration” for the product Optic White Renewal Toothpaste based on results from that clinical study using changes in YIO as the metric (Hogan, 2021; Colgate-Palmolive Company, 2021). A paper has been written based on work from chapter 5 and is currently being reviewed and waiting approval from Colgate to submit to the Journal of Dentistry in hopes to be published.

7 References

1. Abdi, H. and Williams, L.J. 2010. Principal Component Analysis. *WIREs Computational Statistics.* **2**, pp.433-459.
2. Ahmad, I. 2006. *Protocols for Predictable Aesthetic Dental Restorations*. Oxford: Blackwell Munksgaard.
3. Ahn, J. and Lee, Y. 2008. Color distribution of shade guides in the value, chroma, and hue scale. *The Journal of Prosthetic Dentistry*. **100**(1), pp. 18-28.
4. Al-Azzawi, A. 2007. *Light and optics: principles and practices* . Boca Raton, Florida: CRC/Taylor & Francis.
5. Al-Zarea, B. K. 2013. Satisfaction with Appearance and the Desired Treatment to Improve Aesthetics. *International Journal of Dentistry*. **2013**, pp. 1-7.
6. Alkhatib, M. N., Holt, R. and Bedi, R. 2005. Age and perception of dental appearance and tooth colour. *Gerodontology*. **22**, pp. 32-36.
7. Alqahtani, M. Q. 2014. Tooth-bleaching procedures and their controversial effects: A literature Review. *The Saudi Dental Journal*. **26**, pp. 33-46.
8. ASTM E1164-12. 2017. ASTM E1164-12(2017):Standard Practice for Obtaining Spectrometric Data for Object-Color Evaluation. West Conshohocken, PA: ASTM International.
9. ASTM E1347-06. 2020. ASTM E1347-06: Standard Test Method for Color and Color-Difference Measurement by Tristimulus Colorimetry. West Conshohocken, PA: *ASTM International*.
10. ASTM E1349-06. 2018. ASTM E1349-06(2018): Standard Test Method for Reflectance Factor and Color by Spectrophotometry Using Bidirectional (45°:0° or 0°:45°) Geometry. West Conshohocken, PA: ASTM International.
11. ASTM E179-17. 2017. ASTM E179-17: Standard Guide for Selection of Geometric Conditions for Measurement of Reflection and Transmission Properties of Materials. West Conshohocken, PA: ASTM International.

12. ASTM E2466-13. 2013. ASTM E2466-13: Standard Test Method for Colorimetry of Teeth Using Digital Still Camera Technology. West Conshohocken, PA: ASTM International.
13. ASTM E2545-07. 2007. ASTM E2545-07: Standard Test Method for Objective Measurement of Gingival Color Using Digital Still Cameras. West Conshohocken, PA: ASTM International.
14. ASTM E308-18. 2001. ASTM E308-18: Standard Practice for Computing the Colors of Objects by Using the CIE System. West Conshohocken, PA: ASTM International.
15. ASTM E313-05. 2005. ASTM E313-05:Standard Practice for Calculating Yellowness and Whiteness Indices from Instrumentally Measured Color Coordinates. E313-05. West Conshohocken, PA: *ASTM International*.
16. ASTM G138-12. 2012. ASTM G138-12: Standard Test Method for Calibration of a Spectroradiometer Using a Standard Source of Irradiance. West Conshohocken, PA: ASTM International.
17. Bahannan, S.A. 2014. Shade matching quality among dental students using visual and instrumental methods. *Journal of Dentistry*. **42**, pp. 48-52.
18. Baltzer, A. and Kaufmann-Jinoian, V. 2004. The Determination Of The Tooth Colors. *Quintessenz Zahntechnik.* **30**(7), pp. 726-740.
19. Beltrami, R., Colombo, M., Chiesa, M., Bianchi, S. and Poggio, C. 2014. Scattering properties of a composite resin: Influence on color perception. *Contemporary clinical dentistry*, **5**(4), pp.501–506.
20. Bengel, W. 2002. *Mastering Dental Photography.* Berlin, Germany: Quintessenz Verlags-GmbH.
21. Billmeyer, F.W. and Saltzman, M. 1981. *Principles of color technology.* 2nd ed. New York: Wiley.
22. Blair, J.J. 1985. *Dental Restoration Shading*. EP0154137 A1. 1985-09-11.
23. Blum, S., Horn, M. and Olms, C. 2018. A comparison of intraoral spectrophotometers—Are there user-specific differences? *Journal of Esthetic and Restorative Dentistry*. **30**(5), pp. 442–448.
24. Bolt, R. A., Ten Bosch, J. J. and Coops, J. C. 1994. *Influence of window size*

in small-window colour measurement, particularly of teeth. Physics in Medicine and Biology. **39**, pp. 1133-1142.

25. Bona, A. D., Barrett, A. A., Rosa, V. and Pinzetta, C. 2009. Visual and instrumental agreement in dental shade selection: Three distinct observer populations and shade matching protocols. *Dental Materials.* **25**, pp. 276-281.
26. Brandt, J., Nelson, S., Lauer, H., von Hehn, U. and Brandt, S. 2017. In vivo study for tooth colour determination—visual versus digital. *Clinical Oral Investigations*, **21**(9), pp. 2863–2871.
27. Brannan, D., Esplen, M. and Gray, J. 2011. *Geometry.* 2nd ed. Cambridge: Cambridge University Press.
28. Brook, A. H., Smith, R. M. and Lath, D. J. 2007. The clinical measurement of tooth colour and stain. *International Dental Journal*. **57**, pp. 324-330.
29. Burkinshaw, S. 2004. Colour in relation to dentistry. Fundamentals of colour science. *British Dental Journal*. **196**(1), pp. 33–41.
30. Cal, E., Sonugelen, M., Guneri, P., Kesercioglu, A. and Kose, T. 2004. Application of a digital technique in evaluating the reliability of shade guides. *Journal of Oral Rehabilitation*. **31**(5), pp. 483–491.
31. Carey, C. 2014. Tooth Whitening: What we now know. J*ournal of evidence-based dental practice special issue -annual report on dental hygiene.* **14**(1), pp. 70-76.
32. Cho, B., Lim, Y. and Lee, Y. 2007. Comparison of the color of natural teeth measured by a colorimeter and Shade Vision System. *Dental Materials*. **23**(10), pp. 1307–1312.
33. Chu, S.J. 2007. Range And Mean Distribution Frequency Of Individual Tooth Width Of The Maxillary Anterior Dentition. *Practical Procedures In Aesthetic Dentistry*. **19**(4), Pp. 209-215.
34. Chu, S.J. and Okubo, S. 2008. Range and mean discordance of individual tooth width of the mandibular anterior dentition. *Practical Procedures in Aesthetic Dentistry*. **20**(5), pp. 313-320.

35. Chu, S.J., Trushkowsky, R.D. and Paravina, R.D.. 2010. Dental color matching instruments and systems: review of clinical and research aspects. *Journal of dentistry*. **38s**, pp. E2-e16.
36. Church, T., Jessup, J.P., Guillory, V.L and Vandewalle, K.S. 2017. Translucency and strength of high-translucency monolithic zirconium oxide materials. *General Dentistry.* (398), pp. 48-52.
37. CIE 15. 2004. CIE 15: Colorimetry Technical Report. Commission Interntionale de l'Eclairage. Vienna, Austria: CIE Central Bureau.
38. CIE 217. 2016. CIE 217: Recommended Method for Evaluating the Performance of Colour-Difference Formulae Technical Report. Commission Interntionale de l'Eclairage. Vienna, Austria: CIE Central Bureau.
39. Clarke, F. J. and Parry, D. 1985. Helmholtz Reciprocity: Its validity and application to reflectometry. *Lighting Research & Technology (1969)*. **17**(1), 1–11.
40. Colgate-Palmolive Company. 2021. *Optic White Renewal Toothpaste*. [Online]. [7 July 2021]. Available from: https://www.colgate.com/en-us/products/toothpaste/ow-renewal
41. Collins, L.Z., Naeeni, M., and Platten, S.M. 2008. Instant tooth whitening from a silica toothpaste containing blue covarine. *Journal of Dentistry.* **36s**, pp. S21-S25.
42. Curd, F.M., Jasinevicius, T.R., Graves, A., and Cox, V. 2006. Comparison of the shade matching ability of dental students using two light sources. *Journal of Prosthetic Dentistry*. **96**(6), pp. 391-396.
43. Da Silva, J. D., Park, S. E., Weber, H. and Ishikawa-Nagai, S. 2008. Clinical Performance of a newly developed spectrophotometric system on tooth colour reproduction. *The Journal of Prosthetic Dentistry.* **99**(5). pp. 361-368.
44. De Bragança, R.M., Moraes, R.R. and Faria-e-Silva, A.L. 2021. Color assessment of resin composite by using cellphone images compared with a spectrophotometer. *Restorative Dentistry & Endodontics.* **46**, pp. e23.
45. del Mar Pérez, M., Ghinea, R., Rivas, M.J., Yebra, A., Ionescu, A.M., Paravina, R.D. and Herrera, L.J. 2016. Development of a customized whiteness index for dentistry based on CIELAB colour space. *Dental*

Materials. **32**, pp. 461-467.

46. Ebner, M. ed. 2007. *Color Constancy.* West Sussex: John Wiley & Sons, Ltd.
47. Elamin, H.O., Abubakr, N.H. and Ibrahim, Y.E. 2015. Identifying the tooth shade in group of patients using Vita EasyShade. *European Journal of Dentistry.* **9**(2), pp. 213-217.
48. Finlayson, G., and Zhu, Y. 2019. Finding a Colour Filter to Make a Camera Colorimetric by Optimisation. In: *Computational Color Imaging*, 27-29 March 2019, Japan. Switzerland: Springer International Publishing, pp. 53-62.
49. Flucke, J. 2021. The Future of Teeth Whitening. *Dental Products Report.* **55**(10), pp. 22-23.
50. García, P.A., Huertas, R., Melgosa, M. and Cui, G. 2007. Measurement of the relationship between perceived and computed color difference. *Journal of the Optical Science of America A.* **24**(7), pp. 1823-1829.
51. German, D.S., Chu, S.J., Fulong, M.L. and Patel, A. 2016. Simplifying optimal tooth-size calculations and communications between practitioners. *American Journal of Orthodontics and Dentofacial Orthopedics*. **150**(6), pp. 1051-1055.
52. Gevaux, L., Simonot, L., Clerc, R., Gerardin, M. and Hebert, M. 2020. Evaluating edge loss in the reflectance measurement of translucent materials. *Applied Optics. Optical Technology and Biomedical Optics*. **59**(28), pp. 8939–8950.
53. Gómez-Polo, C., Gómez-Polo, M., Celemin-Viñuela, A. and Martínez Vázquez De Parga, J. 2014. Differences between the human eye and the spectrophotometer in the shade matching of tooth colour. *Journal of Dentistry*. **42**(6), pp. 742–745.
54. Gómez-Polo, C., Muñoz, M.P., Lorenzo Luengo, M.C., Vicente, P., Galindo, P. and Martín Casado, A.M. 2015. Comparison of the CIELab and CIEDE2000 color difference formulas. *Journal of Prosthetic Dentistry.* **115**(1), pp. 65-70.

55. Gozalo-Diaz, D.J., Lindsey, D.T., Johnston, W.M. and Wee, A.G. 2007. Measurement of color for craniofacial structures using a 45/0-degree optical configuration. Journal of Prosthetic Dentistry. **97**(1), pp.45-53.
56. Grand View Research. 2021. *Oral Care Market Size, Share and Trends Report.* [Online]. Market Analysis Report: Grand View Research. [Accessed 31 October 2021]. Available from: https://www.grandviewresearch.com/industry-analysis/oral-care-market
57. Guan, Y. H., Lath, D. L., Lilley, T. H., Willmot, D. R., Marlow., I. and Brook, A.H. 2005. The measurement of tooth whiteness by image analysis spectrophotometry: a comparison. *Journal of Oral Rehabilitation.* **32**, pp. 7-15.
58. Gupte, V.C. 2010. Expressing colours numerically. In: Gulrajani, M. L., Colour measurement Principles, advances and industrial applications. Oxford: Woodhead Publishing Limited.
59. Hammad, I. A. 2003. Intrarater repeatability of shade selections with two shade guides. *The Journal of Prosthetic Dentistry.* **89**(1), pp. 50-53.
60. He, L. and Nahar, V. 2016. Reuse of scientific data in academic publications: An investigation of Dryad Digital Repository. *Aslib Journal of Information Management*. **68**(4), pp. 478–494.
61. He, W., Park, C., Byun, S., Tan, D., Lin, C. and Chee, W. 2020. Evaluating the relationship between tooth color and enamel thickness, using twin flash photography, cross-polarization photography, and spectrophotometer. *Journal of Esthetic and Restorative Dentistry*. **32**(1), pp. 91–101.
62. Hess, A.S. and Hess, J.R. 2018. Principal Component Analysis. *Transfusion.* **58**, pp.1580-1582.
63. Hirschen, R. 2010. Visual and Instrumental Evaluation of whiteness and yellowness. In: Gulrajani, M. L., Colour measurement Principles, advances and industrial applications. Oxford: Woodhead Publishing Limited.
64. Höfel, L, Lange, M. and Jacobsen, T. 2007. Beauty and the teeth: perception of tooth color and its influence on the overall judgment of facial

attractiveness. *The International Journal of Periodontics & Restorative Dentistry*. **27**(4), pp. 349–357.

65. Hogan, R. 2021. Email Response to Chelsea Sullivan, 29 January 2021.
66. Howitt, D. and Cramer, D. 2014. *Introduction to Statistics in Psychology.* 6th ed. Harlow, England: Pearson.
67. Hughes, G. 2017. Sharing research data. *Emergency Medicine Australasia.* **29**, pp. 4-5.
68. Hunt, R.W.G and Pointer, M.R. 2011. Measuring Colour. 4th Ed. West Sussex: John Wiley & Songs, Ltd.
69. Hunter, R. S. 1981. Conversion of Visual to Instrumental Measurement of Yellowness. *Journal of the American Oil Chemists' Society.* **59**(5), pp. 608-612
70. Jarad, F., Russell, M. and Moss, B. 2005. The use of digital imaging for colour matching and communication in restorative dentistry. *British Dental Journal*. **199**(1), pp. 43–49.
71. Jeng, Y., Lin, T., Hsu, H., Chang, H. and Shieh, D. 2011. Human enamel rod presents anisotropic nanotribiological properties. *Journal of the mechanical behavior of biomedical materials*. **4**, pp. 515-522.
72. Johnston, W. M. 2009. Color measurement in dentistry. *Journal of Dentistry*. **37s**, pp. e2-e6.
73. Johnston, W.M., Hesse, N.S., Davis, B.K. and Seghi, R.R. 1996. Analysis of Edge-losses in Reflectance Measurements of Pigmented Maxillofacial Elastomer. *Journal of Dental Research.* **75**(2), pp. 752-760.
74. Joiner, A. 2004. Tooth Colour: A Review Of The Literature. *Journal Of Dentistry*. **32**, pp. 3-12
75. Joiner, A. 2010. Whitening toothpastes: A review of the literature. *Journal of Dentistry*. **38s**, pp. e17-e24.
76. Joiner, A. and Luo, W. 2017. Tooth colour and whiteness: A review. *Journal of Dentistry*. **67**, pp. S3-S10.
77. Joiner, A., Hopkinson, I., Deng, Y. and Westland, S. 2008a. A review of tooth colour and whiteness. *Journal of Dentistry*. **36s**, pp. S2-S7.

78. Joiner, A., Philpotts, C.J., Alonso, C., Ashcroft, A.T. and Sygrove, N J. 2008b. A novel optical approach to achieving tooth whitening. *Journal of dentistry*. **36**, pp. 8-14.
79. Kershaw, S., Newton, J. and Williams, D. 2008. The influence of tooth colour on the perceptions of personal characteristics among female dental patients: comparisons of unmodified, decayed and 'whitened' teeth. *British Dental Journal*. **204**(E9).
80. Khashayar, G., Bain, P., Salari, S., Dozic, A., Kleverlaan, C. and Feilzer, A. 2014. Perceptibility and acceptability thresholds for colour differences in dentistry. *Journal of Dentistry*. ***42***(6), pp. 637–644.
81. Khashayar, G., Dozic, A., Kleverlaan, C.J. and Feilzer, A.J. 2012. Data comparison between two dental spectrophotometers. *Operative Dentistry*. **37**(1), pp.12-20.
82. Khurana, R., Tredwin, C.J., Weisbloom, M. and Moles, D. R. 2007. A clinical evaluation of the individual repeatability of three commercially available colour measuring devices. *British Dental Journal*. **203**(12), pp. 3-12.
83. Kim-Pusateri, S., Brewer J. D., Davis, E. L. and Wee, A. G. 2009. Reliability and Accuracy of four dental shade-matching devices. *The Journal of Prosthetic Dentistry.* **101**(3), pp. 193-199.
84. Kim, J. Yu, B. and Lee, Y. 2008. Influence of surface layer removal of shade guide tabs on the measured color by spectrophotometer and spectroradiometer. *Journal of Dentistry.* **36**, pp. 1061-1067.
85. Kim, M., Kim, B., Park, B., Lee, M., Won, Y., Kim, C. and Lee, S. 2018. A Digital Shade-Matching Device for Dental Color Determination Using the Support Vector Machine Algorithm. *Sensors.* **18**(9), pp. 3051.
86. Kim, S.T., Abbott, P.V. and McGinley, P. 2000. The effects of Ledermix paste on discolouration of mature teeth. *International Endodontic Journal*. **33**(3), 227–232.
87. Kim, Y., Son, H., Yi, K., Kim, H., Ahn, J. and Chang, J. 2013. The color change in artificial white spot lesions measured using a spectroradiometer. *Clinical Oral Investigations.* **17**, pp. 139-146.

88. Kingdom, F.A.A. and Prins, N. 2010. Psychophysics: A Practical Introduction. London: Elsevier Ltd
89. Kirchner, E. and Dekker, N. 2011. Performance measures of color-difference equations: correlation coefficient versus standardized residual sum of squares. *Journal of the Optical Science of America A.* **28**(9), pp. 1841-1848.
90. Klotz, A., Habibi, Y., Corcodel, N., Rammelsberg, P., Hassel, A.J. and Zenthöfer, A. 2018. Laboratory and clinical reliability of two spectrophotometers. *Journal of Esthetic and Restorative Dentistry. Pp., 1-5.*
91. Knight, R.D. 2008. *Physics for Scientists and Engineers: A strategic Approach.* 2nd ed. Upper Saddle River, NJ: Pearson.
92. Konica Minolta Sensing Americas Inc., 2020. Color Measurement Instrument Geometries. Konica Minolta Sensing America. [Online]. [Accessed 10 August 2020]. Available from: https://sensing.konicaminolta.us/us/blog/color-measurement-instrument-geometries/
93. Konica Minolta. 2007. Spectroradiometer CS-2000/CS-2000A. [Instruction Manual]. United States: Konica Minolta, Inc.
94. Konica Minolta. 2014. Spectrophotometer CM-2600d/2500d. [Instruction Manual]. Japan: Konica Minolta, Inc.
95. Konica Minolta. 2015. Precise color communication; color control from perception to instrumentation. [Booklet]. Japan: Konica Minolta, Inc.
96. Konica Minolta. 2016a. Spectrophotometer CM-25cG. [Instruction Manual]. United Kingdom: Konica Minolta, Inc.
97. Konica Minolta. 2016b. Spectrophotometer CM-M6. [Instruction Manual]. United States: Konica Minolta, Inc.
98. Konica Minolta. 2019. Specular Component Included (SCI) vs. Specular Component Excluded (SCE). 24 January. Colour Measurement. [Online]. [Accessed 27 July 2020]. Available from: https://sensing.konicaminolta.asia/specular-component-included-sci-vs-specular-component-excluded-sce/

99. Koumpia, E., Athanasiou, A.E., Eliades, T. and Knösel, M. 2018. Precision of a reflectance spectrophotometer in measuring anterior tooth colour. *The Open Dental Journal.* **12***, pp. 884-895.*

100. Kovacevic Pavicic, D., Pavlic, A., Kinkela Devcic, M., Lajnert, V. and Spalj, S. 2017. Tooth Color as a Predictor of Oral Health-Related Quality of Life in Young Adults. *Journal of Prosthodontics*. **28**(4), pp. e886–e892.

101. Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Pychometrika*. **29**, pp. 1–27.

102. Kuehni, R. G. 2003. *Color Space and Its Divisions: Color Order from Antiquity to the present.* Hoboken, New Jersey: John Wiley & Sons Inc.

103. Langlais, R.P., Miller, C.S. and Nield-gehrig, J.S. 2009. Color Atlas of Common Oral Diseases. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins.

104. Lasseree, J., Pop-Ciutrila, I. and Colosi, H. 2011. A comparison between a new visual method of colour matching by intraoral camera and conventional visual and spectrometric methods. *Journal of Dentistry*. **39s**, pp. e29-e36.

105. Lee, Y. and Powers, J.M. 2005. Comparison of CIE Lab, CIEDE 2000, and DIN 99 Color Differences Between Various Shade of Resin Composites. *The International Journal of Prosthodontics.* **18**(2), pp. 150-155.

106. Lee, Y.K., Lim, B.S., and Kim, C.W. 2004. Influence of illuminating and viewing aperture size on the color of dental resin composites. *Dental Materials*. **20**(2), pp. 116–123.

107. Lee, Y.K., Yu, B., Lim, J. I. and Lim, H. 2011. Perceived color shift of a shade guide according to the change of illuminant. *The Journal of Prosthetic Dentistry.* **105**(2), pp. 91-99.

108. Lehmann, K.M., Igiel, C., Schmidtmann, I. and Scheller, H. 2010. Four color-measuring devices compared with a spectrophotometric reference system. *Journal of Dentistry.* **38s**, pp. e65-e70.

109. Lim, H., Yu, B., Lim, J. and Lee, Y. 2010. Correlations between spectroradiometric and spectrophotometric colors of all-ceramic materials. *Dental Materials*. **26**(11), pp. 1052–1058.

110. Luo, M., Cui, G. and Rigg, B. 2001. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Application*. **26**(5), pp 340–350.

111. Luo, R. M. 2011. The quality of light sources. *Coloration Technology*. **127**, pp. 85-87.

112. Luo, W., Naeeni, M., Platten, S., Wang, J., Sun, J.N., Westland, S. and Joiner, A. 2017. The in vitro and in vivo reproducibility of video-based digital imaging system for tooth colour measurement. *Journal of Dentistry.* **67**, pp. S15-S19.

113. Luo, W., Westland, S., Brunton, P., Ellwood, R., Pretty, I. A. and Mohan, N. 2007. Comparison of the ability of different colour indices to assess changes in tooth whiteness. *Journal of Dentistry.* **35**, pp. 109-116.

114. Luo, W., Westland, S., Ellwood, R., Pretty, I. and Cheung, V. 2009. Development of a whiteness index for dentistry. *Journal of Dentistry.* **37s**, pp. e21-e26.

115. MacAdam, D.L. 1942. Visual Sensitivity to Color Differences in Daylight. *Journal of the Optical Socieiety of America.* **32**(5), pp. 247-274.

116. MacDonald, R. 1980. Industrial pass/fail color matching. Part I. Preparation of visual color-matching data. *Journal of Society of Dyers and Colorists*. **96**, Pp. 372–376.

117. Mackie, I.C. and Blinkhorn, A. 1995. Oral Health Prevention of Dental Disease. Manchester, United Kingdom: Oral Health and Development University Dental Hospital of Manchester.

118. Mahn, E., Tortora, S., Olate, B., Cacciuttolo, F., Kernitsky, J., and Jorquera, G. 2021. Comparison of visual analog shade matching, a digital visual method with a cross-polarized light filter, and a spectrophotometer for dental color matching. *The Journal of Prosthetic Dentistry*. **125**(3), pp. 511–516.

119. Malacara, D. 2002. Color Vision and Colorimetry: theory and applications. Washington: SPIE Press.

120. Marques Junior, R.B., Baroudi, K., Campos dos Santos, A.F., Pontes, D. and Amaral, M. 2021. Tooth Discoloration Using Calcium Silicate-Based

Cements For Simulated Revascularization in Vitro. *Brazilian Dental Journal*. **32**(1), pp. 53-58.

121. Martínez-Verdú, F., Chorro, E., Parales, E., Vilaseca, M. and Pujol, J. 2010. Camera-based colour measurement. In: Gulrajani, M. L., Colour measurement Principles, advances and industrial applications. Oxford: Woodhead Publishing Limited.
122. Mazur, M., Westland, S., Jedliński, M., Maruotti, A., Nardi, G., Ottolenghi, L. and Guerra, F. 2020. The Influence of Dental Occlusion on Spectrophotometric Tooth Color Determinations. *The Open Dentistry Journal.* **14**, pp. 247-254.
123. McDonald, R. 1997. *Colour Physics for Industry.* 2nd ed. Bradford: Society of Dyers and Colourists.
124. McLaren, K. 1970. Colour Passing - Visual or Instrumental. *Journal of Society of Dyers and Colourists*. **86**(9), pp. 389-392
125. Melgosa, M., Alman, D., Grosman, M., Gómez-Robledo, L., Trémeau, A., Cui, G., García, P., Vázquez, D., Li, C. and Luo, M. 2013. Practical demonstration of the CIEDE2000 corrections to CIELAB using a small set of sample pairs. *Color Research and Application*. **38**(6), 429–436.
126. Melgosa, M., Huertas, R. and Berns, R.S. 2008. Performance of recent advanced color-difference formulas using the standardized residual sum of squares index. *Journal of the Optical Science of America A.* **25**(7), pp. 1828-1834.
127. Mello, M.M., Lieou, V. and Goodman, S.N. 2018. Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing. *The New England Journal of Medicine.* **378**, pp. 2202-2211.
128. Mishchenko, M.I., Travis, L.D. and Lacis, A.A. 2002. *Scattering, absorption, and emission of light by small particles*. Bambridge: Cambridge University Press.
129. Mohan, N., Westland, S., Brunton, P., Ellwood, R., Pretty, I.A. and Wen, L. 2008. A clinical study to evaluate the efficacy of a novel tray based tooth whitening system. *Journal of Dentistry*. **26**, pp. 21-26.

130. Montero, J., Gomez-Polo, C., Santos, J. A., Portillo, M., Lorenzo, M. C. and Albaladejo, A. 2014. Contributions of dental color to the physical attractiveness stereotype. *Journal of Oral Rehabilitation*. **41**, pp. 768-782.

131. Nassau, K. 1998. *Color for science, art and technology*. Amsterdam, Netherlands: Elsevier.

132. National Library of Medicine. 2021. *Data Reuse*. [Online]. [Accessed 1 February 2021]. Available from: https://nnlm.gov/data/thesaurus/data-reuse

133. Nedelcu, R., Olsson, P., Nyström, I., Rydén, J. and Thor, A. 2018. Accuracy and precision of 3 intraoral scanners and accuracy of conventional impressions: A novel in vivo analysis method. *Journal of Dentistry*. **69**, pp. 110-118.

134. O'Brien, W.J., Groh, C.L. and Boenke, K.M. 1990. A new, small-color-difference equation for dental shades. *Journal of Dental Research.* **69**(11), pp. 1762-1764.

135. Odioso L.L., Gibb, R.D. and Gerlach, R.W. 2000. Impact of demographic, behavioural, and dental care utilization parameters on tooth color and personal satisfaction. *Compendium of Continuing Education in Dentistry.* **21**(29), pp. S35—S41.

136. Oguro, R., Nakajima, M., Seki, N., Sadr, A., Tagami, J. and Sumi, Y. 2016. The role of enamel thickness and refractive index on human tooth colour. *Journal of Dentistry*. **51**, pp. 36-44.

137. Oleari, C. 2016. *Standard Colorimetry: Definitions, Algorithms and Sotware.* United Kingdom: John Wiley & Sons, Ltd.

138. Oliveira, D., Ayres, A., Rocha, M., Giannini, M., Puppin Rontani, R., Ferracane, J. and Sinhoreti, M. 2015. Effect of Different In Vitro Aging Methods on Color Stability of a Dental Resin-Based Composite Using CIELAB and CIEDE2000 Color-Difference Formulas. *Journal of Esthetic and Restorative Dentistry*. ***27***(5), pp. 322–330

139. Palumbo, P.A. and Weber, W.L. 1999. Novel spectrophotometer for the measurement of colour and appearance. *Analytica Chimica Acta*. **380**(2), 243-261.

140. Pan, Q. and Westland, S. 2018. Tooth colour and whitening – digital

technologies. *Journal of Dentistry.* **74**(1), pp. S42-S46.

141. Pan, Q., Westland, S. and Ellwood, R. 2018. Evaluation of colorimetric indices for the assessment of tooth whiteness. *Journal of Dentistry.* **76**, pp. 132-136.

142. Parameswaran, V., Anilkumar, S., Lylajam, S., Rajesh, C. and Narayan, V. 2016. Comparison of accuracies of an intraoral spectrophotometer and conventional visual method for shade matching using two shade guide systems. *The Journal of Indian Prosthodontic Society*. **16**(4), pp. 352–358.

143. Paravina, R. D. 2008. *New Shade Guide for Tooth Whitening Monitoring: Visual Assessment.* The Journal of Prosthetic Dentistry. **99**(3), pp. 178-184.

144. Paravina, R., Ghinea, R., Herrera, L., Bona, A., Igiel, C., Linninger, M., Sakai, M., Takahashi, H., Tashkandi, E. and Del Mar Perez, M. 2015. Color Difference Thresholds in Dentistry: Color Difference Thresholds. *Journal of Esthetic and Restorative Dentistry*. **27**, pp. S1–S9.

145. Paravina, R.D. and Powers, J. 2004. *Esthetic color training in dentistry*. St. Louise, MO: Elsevier Mosby.

146. Paravina, R.D., Aleksić, A., Tango, R.N., García-Beltrán, Johnston, W.M. and Ghinea, R.I. 2021. Harmonization of color measurements in dentistry. *Measurement.* **169**, pp. 108504

147. Paravina, R.D., O'Keefe, K.L. and Kuljic, B.L. 2006. Color of permanent teeth: a prospective clinical study. *Balkan Journal of Stomatology*. **10**, pp. 93–97.

148. Pasquetto, I.V., Sands, A.E. and Borgman, C. L. 2015. Exploring Openness in Data and Science: What is "Open," to Whom, When, and Why?. Proceedings of the Association for Information Science and Technology. **52**(1), pp. 1-2.

149. Paul, S., Peter, A., Pietrobon, N. and Hämmerle, C. 2002. Visual and Spectrophotometric Shade Analysis of Human Teeth. *Journal of Dental Research*. **81**(8), pp. 578–582.

150. Pérez, M., Ghinea, R., Rivas, M.J., Yebra, A., Ionescu, A.M., Paravina, R.D. and Herrera, L.J. 2016. Development of a customized whiteness index

for dentistry based on CIELAB colour space. *Dental Materials.* **32**, pp. 461-467.

151. Piwowar, H.A. and Vision, T.J. 2013. Data reuse and the open data citation advantage. *PeerJ.* **1,** pp. e175.

152. Pop-Ciutrila, I. S., Ghinea, R., Colori, H. A. and Dunea, D. 2016. Dentin translucency and color evaluation in human incisors, canines, and molars. *Journal of Prosthetic Dentistry.* **115**(4), pp. 475-481.

153. Purdue, 2015. *F Critical Values Table.* [Online]. [Accessed 20 September 2018]. Available from: https://www.stat.purdue.edu/~jtroisi/STAT350Spring2015/tables/FTable.pdf

154. Rasmussen, C.M. 2019. Market Analysis 2020 on 26th World Congress on Dentistry and Oral Health. *Journal of Oral Medicine***. 3**(5).

155. Rauf, W. 2020. Email response to Chelsea Sullivan, 20 July 2020.

156. Rubio, L., Sioli, J.M., Suarez, J., Gaitan, M. J. and Martin-de-las-Heras, S. 2015. Spectrophotometric analysis of color changes in teeth incinerated at increasing temperatures. *Forensic Science International*. **252**, pp. 193.e1–193.e6.

157. Schmidseder, J. 2000. Aesthetic Denistry. New York: Thieme Stuttgart.

158. Seghi, R.R. 1990. Effects of Instrument-measuring geometry on Colorimetric Assessments of Dental Porcelains. *Journal of Dental Restorations. **69***(5), pp. 1180-1183.

159. Sindi, K.H., Bubb, N.L., Gutteridge, D.L. and Evans, J.A. 2015. In vitro enamel thickness measurements with ultrasound. *Ultrasound in Medicine and Biology.* **41**(1), pp. 301-308.

160. Stratview Research, 2020. *Global Teeth Whitening Products Market, Dynamics, and Market Analysis.* Radian Offshore Consultanacy LLP.

161. Sullivan, C., Pan, Q., Westland, S. and Ellwood, R. 2019. A yellowness index for use in dentistry. *Journal of Dentistry.* **91**, pp. 103244.

162. Tao, D., Smith, R.N., Zhang, Q., Sun, J.N., Philpotts, C.J., Ricketts, S.R., Naeeni, M. and Joiner, A. 2017a. Tooth whitening evaluation of blue covarine containing toothpastes. *Journal of Dentistry*. 67, pp. S20-S24.

163. Tao, D., Sun, J.N., Wang, X., Zhang, Q., Naeeni, M.A., Phipotts, C.J. and Joiner, A. 2017b. In vitro and clinical evaluation of optical tooth whitening toothpastes. Journal of Dentistry. **67**, pp. S25–S28.

164. Ten Bosch J.J. and Coops J.C. 1995. Tooth color and reflectance as related to light scattering and enamel hardness. *Journal of Dental Research.* **74**(1), pp. 74-80.

165. Thurstone, L.L. 1927. Psychophysical Analysis. *The American Journal of Psychology.* **38**(3), pp. 368-389.

166. Tin-Oo, M., Saddki, N. and Hassan, N. 2011. Factors influencing patient satisfaction with dental appearance and treatments they desire to improve aesthetics. *BMC Oral Health*. **11**(6).

167. Tsukida, K. and Gupta, M.R. 2011. How to Analyze Paired Comparison Data. UWEETR-2011-0004 Electrical Engineering. May 2011, University of Washington Department of Electrical Engineering.

168. Tung, F. F., Goldstein, G. R., Jang, S. and Hittelman, E. 2002. The repeatability of an intraoral dental colorimeter. *The Journal of Prosthetic Dentistry.* **88**(6), pp. 585-590.

169. Vallittu, P. K., Vallittu, A. S. and Lassila, V. P. 1996. Dental aesthetics – a survey of attitudes in different groups of patients. *Journal of Dentistry*. **24**(5), pp. 335-338.

170. VeriVide. 2008. *Powerful, fast and flexible colour management: DigiEye can evaluate objects previously considered unmeasurable.* [Leaflet]. Leicester: Verivide.

171. VeriVide. 2020. DigiEye System: Non-contact digital colour measurement and imaging system. [Product Specification Leaflet]. Leicester, United Kingdom: VeriVide.

172. VITA North America. 2022. *VITA classical Shade Guide*. [Online]. [Accessed 22 May 2022]. Available from: https://www2.vitanorthamerica.com/products/shade-management/vita-classical-previously-the-lumin-vacuum-shade-guide/

173. VITA Zahnfabrik. 2013. *VITA Easyshade Advance 4.0*. [Operating Instructions]. Germany: VITA Zahnfabrik H. Rauter GmbH & Co.KG.

174. VITA Zahnfabrik. 2020a. VITA Toothguide 3D-MASTER. [Instruction for use]. Germany: VITA Zahnfabrik H. Rauter GmbH & Co.KG.

175. VITA. 2020b. VITA EasyShade Advance 4.0. [Online]. [3 September 2020]. Available from: https://www.vita-zahnfabrik.com/en/VITA-Easyshade-Advance-40-17073.html

176. Wang, F., Takahashi, H. and Iwasaki, N. 2013. *Translucency of Dental Ceramics with Different Thickness.* The Journal of Prosthetic Dentistry. **110**(1), pp. 14-20.

177. Wee, A. G., Lindsey, D. T., Kuo, S. and Johnston, W.M. 2006. Color accuracy of commercial digital cameras for use in dentistry. *Dental Materials*. **22**. pp. 553-559.

178. Westland, S., Luo, W., Li, Y., Pan, Q. and Joiner, A. 2017. Investigation of the perceptual thresholds of tooth whiteness. *Journal of Dentistry.* **67**, pp. S11-S14.

179. Westland, S., Ripamonti, C. and Cheung, V. 2012. *Computational Colour Science using MATLAB.* West Sussex: John Wiley & Sons, Ltd.

180. Wyszecki, G. and Stiles, W.S. 1967. *Color Science Concept and Methods, Quantitative Data and Formulas. 1st ed.* New York: John Wiley & Sons, Inc.

181. X-rite. 2021.962/964 Spectrophotometer Operator's Manual. [Instruction Manual]. Michigan, United States: X-Rite, Inc.

182. Xiao, J., Zhou, X., Zhu, W., Zhang, B., Li, J. and Xu, X. 2007. The prevalence of tooth discolouration and the self-satisfaction with tooth colour in a Chinese urban population. *Journal of Oral Rehabilitation*. **34**(5), pp. 351–360.

183. Yamanel, K., Caglar, A., Özcan, M., Gulsah, K. and Bagis, B. 2010. Assessment of Color Parameters of Composite Resin Shade Guides Using Digital Imaging versus Colorimeter. *Journal of Esthetic and Restorative Dentistry*. **22**(6), 379–388.

184. Yu, B., Ahn, J. and Lee, Y. 2009. Measurement of translucency of tooth

enamel and dentin. *Acta Odontologica Scandinavica.* **67**, pp. 57-64.

185. Yuan, J.C., Brewer, J.D., Monaco, E.A. and Davis, E.L. 2007. Defining a natural tooth color space based on a 3-dimensional shade system. *Journal of Prosthetic Dentistry*. **98**, pp. 110–119.

186. Zenthöfer, A., Cabrera, T., Corcodel, N., Rammelsberg, P. and Hassel, A.J. 2014. Comparison of the Easyshade compact and advance in vitro and in vivo. *Clinical Oral Investigations*. **18**(5), pp.1473-1479

187. Zhang, B., Nie, W. and Zhao, S. 2018. A novel Color Rendition Chart for digital tongue image calibration. *Color Research and Application.* **43**, pp. 749-759.

188. Zhang, X., Wang, Q., Li, J., Zhou, X., Yang, Y. and Xu, H. 2017. Estimating spectral reflectance from camera responses based on CIE XYZ tristimulus values under multi-illuminants. *Color Research and Application.* ***42***(1), 68–77.

Appendix A

MATLAB CODE r2xyz

```
% ================================================
% *** FUNCTION r2xyz
% *** function [xyz] = r2xyz(p, startw, endw, obs)
% *** computes XYZ from reflectance p
% *** p is an n by w matrix of n spectra
% *** e.g. set obs to 'd65_64 for D65 and 1964
% *** the startw and endw variables denote first and
% *** last wavelengths (e.g. 400 and 700) for p which
% *** must be 10-nm data in the range 360-780
% ================================================
function [xyz] = r2xyz(p, startw, endw, obs)

if ((endw>780) | (startw<360) | (rem(endw,10)~=0) | (rem(startw,10)~=0))
   disp('startw and endw must be divisible by 10')
   disp('wavelength range must be 360-780 or less');
   return;
end

load weights.mat
% weights.mat contains the tables of weights
if strcmp('a_64',obs)
   cie = a_64;
elseif strcmp('a_31', obs)
   cie = a_31;
elseif strcmp('c_64', obs)
   cie = c_64;
elseif strcmp('c_31', obs)
   cie = c_31;
elseif strcmp('d50_64', obs)
   cie = d50_64;
elseif strcmp('d50_31', obs)
   cie = d50_31;
elseif strcmp('d55_64', obs)
   cie = d55_64;
elseif strcmp('d55_31', obs)
   cie = d55_31;
elseif strcmp('d65_64', obs)
   cie = d65_64;
elseif strcmp('d65_31', obs)
   cie = d65_31;
elseif strcmp('d75_64', obs)
   cie = d75_64;
elseif strcmp('d75_31', obs)
   cie = d75_31;
elseif strcmp('f2_64', obs)
   cie = f2_64;
elseif strcmp('f2_31', obs)
   cie = f2_31;
elseif strcmp('f7_64', obs)
   cie = f7_64;
elseif strcmp('f7_31', obs)
   cie = f7_31;
elseif strcmp('f11_64', obs)
```

```
   cie = f11_64;
elseif strcmp('f11_31', obs)
   cie = f11_31;
else
   disp('unknown option obs');
   disp('use d65_64 for D65 and 1964 observer');
   return;
end

% check dimensions of P
dim = size(p);
N = ((endw-startw)/10 + 1);
if (dim(2) ~= N)
   disp('dimensions of p inconsistent');
   return;
end

% deal with possible truncation of reflectance
i = (startw - 360)/10 + 1;
if (i>1)
   cie(i,:) = cie(i,:) + sum(cie(1:i-1,:));
end
j = i + N - 1;
if (j<43)
   cie(j,:) = cie(j,:) + sum(cie(j+1:43,:));
end
cie = cie(i:j,:);

% the main calculation
xyz = (p*cie)*100/sum(cie(:,2));

end
% ================================================
% *** END FUNCTION r2xyz
% ================================================
```

Appendix B

MATLAB r2xyz modification code for spectroradiometer (1nm interval) data

```
% ================================================
% *** FUNCTION r2xyz_ Modified 2 for 1nm interval data
% *** and for 2 Degree Observer
% *** function [xyz] = r2xyz(p, startw, endw, obs)
% *** computes XYZ from reflectance p
% *** p is an n by w matrix of n spectra
% *** e.g. set obs to 'd65_31 for D65 and 1931
% *** the startw and endw variables denote first and
% *** last wavelengths (e.g. 400 and 700) for p which
% *** must be 10-nm data in the range 360-780
% ================================================
% modified to work with 1nm data
function [xyz] = r2xyz_mod2(p, startw, endw, obs)
if ((endw>780) | (startw<360) | (rem(endw,1)~=0) | (rem(startw,1)~=0))
   disp('startw and endw must be divisible by 1')
   disp('wavelength range must be 360-780 or less');
   return;
end
load weights_mod2.mat
% weights_mod.mat contains the tables of weights
% whatever weights are in here need to have been interpolated to 1nm
if strcmp('d65_31',obs)
   cie = d65_31;
else
   disp('unknown option obs');
   disp('weight only available for d65_31 (D65 and 1931 observer)');
   return;
end
% check dimensions of P
dim = size(p);
N = ((endw-startw) + 1);
if (dim(2) ~= N)
   disp('dimensions of p inconsistent');
   return;
end
% deal with possible truncation of reflectance
i = (startw - 360) + 1;
if (i>1)
   cie(i,:) = cie(i,:) + sum(cie(1:i-1,:));
end
j = i + N - 1;
if (j<421)
   cie(j,:) = cie(j,:) + sum(cie(j+1:421,:));
end
cie = cie(i:j,:);
% the main calculation
xyz = (p*cie)*100/sum(cie(:,2));
end
% ================================================
% *** END FUNCTION r2xyz
% ================================================
```

Appendix C

MATLAB code xyz2lab

```
% ================================================
% *** FUNCTION xyz2lab
% ***
% *** function [lab] = xyz2lab(xyz, obs, xyzw)
% *** computes LAB from XYZ
% *** xyz is an n by 3 matrix
% *** e.g. set obs to 'd65_64 for D65 and 1964
% *** set obs to 'user' to use optional argument
% *** xyzw as the white point
% ================================================

function [lab] = xyz2lab(xyz,obs,xyzw)

if (size(xyz,2)~=3)
   disp('xyz must be n by 3'); return;
end
lab = zeros(size(xyz,1),size(xyz,2));

if strcmp('a_64',obs)
    white=[111.144 100.00 35.200];
elseif strcmp('a_31', obs)
    white=[109.850 100.00 35.585];
elseif strcmp('c_64', obs)
    white=[97.285 100.00 116.145];
elseif strcmp('c_31', obs)
    white=[98.074 100.00 118.232];
elseif strcmp('d50_64', obs)
    white=[96.720 100.00 81.427];
elseif strcmp('d50_31', obs)
    white=[96.422 100.00 82.521];
elseif strcmp('d55_64', obs)
    white=[95.799 100.00 90.926];
elseif strcmp('d55_31', obs)
    white=[95.682 100.00 92.149];
elseif strcmp('d65_64', obs)
    white=[94.811 100.00 107.304];
elseif strcmp('d65_31', obs)
    white=[95.047 100.00 108.883];
elseif strcmp('d75_64', obs)
    white=[94.416 100.00 120.641];
elseif strcmp('d75_31', obs)
    white=[94.072 100.00 122.638];
elseif strcmp('f2_64', obs)
    white=[103.279 100.00 69.027];
elseif strcmp('f2_31', obs)
    white=[99.186 100.00 67.393];
elseif strcmp('f7_64', obs)
    white=[95.792 100.00 107.686];
elseif strcmp('f7_31', obs)
    white=[95.041 100.00 108.747];
elseif strcmp('f11_64', obs)
    white=[103.863 100.00 65.607];
elseif strcmp('f11_31', obs)
```

```
    white=[100.962 100.00 64.350];
elseif strcmp('user', obs)
    white=xyzw;
else
   disp('unknown option obs');
   disp('use d65_64 for D65 and 1964 observer'); return;
end

lab = zeros(size(xyz,1),3);

fx = zeros(size(xyz,1),3);
for i=1:3
    index = (xyz(:,i)/white(i) > (6/29)^3);
    fx(:,i) = fx(:,i) + index.*(xyz(:,i)/white(i)).^(1/3);
    fx(:,i) = fx(:,i) + (1-index).*((841/108)*(xyz(:,i)/white(i)) + 4/29);
end

lab(:,1)=116*fx(:,2)-16;
lab(:,2) = 500*(fx(:,1)-fx(:,2));
lab(:,3) = 200*(fx(:,2)-fx(:,3));

end
% ================================================
% *** END FUNCTION xyz2lab
% ================================================
```

Appendix D

ΔE_{ab*} MATLAB code

```
% *** FUNCTION cielabde
% *** function [de, dl, dc, dh] = cielabde(lab1, lab2)
% *** computes colour difference from CIELAB values
% *** using CIELAB formula
% *** inputs must be n by 3 matrices
% *** and contain L*, a* and b* values
% *** see also cmcde, cie94de, and cie00de
function [de,dl,dc,dh] = cielabde(lab1,lab2)
if (size(lab1,1)~=size(lab2,1))
   disp('inputs must be the same size'); return;
end
if (size(lab1,2)~=3 | size(lab2,2)~=3)
   disp('inputs must be n by 3'); return;
end
de = zeros(1,size(lab1,2));
dl = zeros(1,size(lab1,2));
dc = zeros(1,size(lab1,2));
dh = zeros(1,size(lab1,2));
dl = lab2(:,1)-lab1(:,1);
dc = (lab2(:,2).^2 + lab2(:,3).^2).^0.5-(lab1(:,2).^2 + lab1(:,3).^2).^0.5;
dh = ((lab2(:,2)-lab1(:,2)).^2 + (lab2(:,3)-lab1(:,3)).^2 - dc.^2);
dh = (abs(dh)).^0.5;
% get the polarity of the dh term
dh = dh.*dhpolarity(lab1,lab2);
de = (dl.^2 + dc.^2 + dh.^2).^0.5;
end
function [p] = dhpolarity(lab1,lab2)
% function [p] = dhpolarity(lab1,lab2)
% computes polarity of hue difference
% p = +1 if the hue of lab2 is anticlockwise
% from lab1 and p = -1 otherwise
[h1,c1] = cart2pol(lab1(:,2), lab1(:,3));
[h2,c2] = cart2pol(lab2(:,2), lab2(:,3));
h1 = h1*180/pi;
h2 = h2*180/pi;

index = (h1<0);
h1 = (1-index).*h1 + index.*(h1+360);
index = (h2<0);
h2 = (1-index).*h2 + index.*(h2+360);

index = (h1>180);
h1 = (1-index).*h1 + index.*(h1-180);
h2 = (1-index).*h2 + index.*(h2-180);
p = (h2-h1);

index = (p==0);
p = (1-index).*p + index*1;
index = (p>180);
p = (1-index).*p + index.*(p-360);
p = p./abs(p);
end
```

Appendix E

ΔE_{00} MATLAB Code

```
% ================================================
% *** FUNCTION cie00de
% ***
% *** function [de,dl,dc,dh] = cie00de(lab1,lab2,sl,sc,sh)
% *** computes colour difference from CIELAB values
% *** using CIEDE2000 formula
% *** inputs must be n by 3 matrices
% *** and contain L*, a* and b* values
% *** see also cielabde, cmcde, and cie94de
function [de,dl,dc,dh] = cie00de(lab1,lab2,sl,sc,sh)

if (size(lab1,1)~=size(lab2,1))
   disp('inputs must be the same size'); return;
end

if (size(lab1,2)~=3 | size(lab2,2)~=3)
   disp('inputs must be n by 3'); return;
end
if (nargin<5)
   disp('using default values of l:c')
   sl=1; sc=1; sh=1;
end

de = zeros(1,size(lab1,2));
dl = zeros(1,size(lab1,2));
dc = zeros(1,size(lab1,2));
dh = zeros(1,size(lab1,2));

% convert the cartesian a*b* to polar chroma and hue
[h1,c1] = cart2pol(lab1(:,2), lab1(:,3));
[h2,c2] = cart2pol(lab2(:,2), lab2(:,3));
h1 = h1*180/pi;
h2 = h2*180/pi;
meanC = (c2+c1)/2;

% compute G factor using the arithmetic mean chroma
G = 0.5 - 0.5*(((meanC.^7)./(meanC.^7 + 25^7)).^0.5);

% transform the a* values
lab1(:,2) = (1 + G).*lab1(:,2);
lab2(:,2) = (1 + G).*lab2(:,2);

% recompute the polar coordinates using the new a*
[h1,c1] = cart2pol(lab1(:,2), lab1(:,3));
[h2,c2] = cart2pol(lab2(:,2), lab2(:,3));
h1 = h1*180/pi;
h2 = h2*180/pi;

% compute the mean values for use later
meanC = (c2+c1)/2;
meanL = (lab2(:,1)+lab1(:,1))/2;
```

```
meanH = (h1+h2)/2;
% Identify positions for which abs hue diff exceeds 180 degrees
meanH = meanH - (abs(h1-h2)>180)*180;
% rollover ones that come -ve
meanH = meanH + (meanH < 0)*360;
% Check if one of the chroma values is zero, in which case set
% mean hue to the sum which is equivalent to other value
index = find(c1.*c2 == 0);
meanH(index) = h1(index)+h2(index);

% compute the basic delta values
dh = (h2-h1);
index = dh>180;
dh = (index).*(dh-360) + (1-index).*dh;
dh = 2*((c1.*c2).^0.5).*sin((dh/2)*pi/180);
dl = lab2(:,1)-lab1(:,1);
dc = c2-c1;

T = 1 - 0.17*cos((meanH-30)*pi/180) + 0.24*cos((2*meanH)*pi/180);
T = T + 0.32*cos((3*meanH + 6)*pi/180) - 0.20*cos((4*meanH - 63)*pi/180);

dthe = 30*exp(-((meanH-275)/25).^2);
rc = 2*((meanC.^7)./(meanC.^7 + 25^7)).^0.5;
rt = -sin(2*dthe*pi/180).*rc;

Lweight = 1 + (0.015*(meanL-50).^2)./((20 + (meanL-50).^2).^0.5);
Cweight = 1 + 0.045*meanC;
Hweight = 1 + 0.015*meanC.*T;

dl = dl./(Lweight*sl);
dc = dc./(Cweight*sc);
dh = dh./(Hweight*sh);

%disp([G T Lweight Cweight Hweight rt])
de = sqrt(dl.^2 + dc.^2 + dh.^2 + rt.*dc.*dh);
```

Appendix F

MATLAB STRESS code

```
% function [pf,fcv] = STRESS(de,dv,f)
% computes performance factor between two datasets: de and dv
% de and dv are the data vectors tested, for example,
% computed colour difference and visual difference
% f is a flag,  f = 1, fcv=1;
%               f != 1, fcv is computed from de,dv
% The stress is a statistical value
% fcv is the scaling factors
% see also none
 function [pf,fcv] = STRESS(de,dv,f)

dim1 = size(de);
dim2 = size(dv);
if (dim1(1) ~= dim2(1)) | (dim1(2) ~= dim2(2))
   disp('the two vectors are not in same size');
   return;
end
if(f==1)
    fcv=1;
else
    fcv=sum(de.*dv)/sum(de.*de);
end

pf = sqrt(sum((fcv*de-dv).^2)/sum(dv.*dv))*100;
```

Appendix G

Ethics Approval

Light Touch Ethical Review Approval from Arts, Humanities and Cultures Faculty Research Ethics Committee at the University of Leeds

The Secretariat
University of Leeds
Leeds, LS2 9JT Tel: 0113 343 4873
Email: ResearchEthics@leeds.ac.uk

UNIVERSITY OF LEEDS

Chelsea Sullivan
School of Design
University of Leeds
Leeds, LS2 9JT

Arts, Humanities and Cultures Faculty Research Ethics Committee
University of Leeds

14 June 2022

Dear Chelsea,

Title of study: **Measurement of Translucent Materials**
Ethics reference: **LTDESN-084**
Grant reference: **113533**

I am pleased to inform you that the above application for light touch ethical review has been reviewed by a representative of the Arts, Humanities and Cultures Faculty Research Ethics Committee and I can confirm a favourable ethical opinion as of the date of this letter. The following documentation was considered:

Document	Version	Date
LTDESN-084 LightTouchEthicsForm.docx	1	22/03/2018

Please notify the committee if you intend to make any amendments to the original research as submitted at date of this approval, including changes to recruitment methodology. All changes must receive ethical approval prior to implementation. The amendment form is available at http://ris.leeds.ac.uk/EthicsAmendment.

Please note: You are expected to keep a record of all your approved documentation, as well as other documents relating to the study. You will be given a two week notice period if your project is to be audited, there is a checklist listing examples of documents to be kept which is available at http://ris.leeds.ac.uk/EthicsAudits.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to ResearchEthics@leeds.ac.uk.

Yours sincerely

Jennifer Blaikie
Senior Research Ethics Administrator, the Secretariat
On behalf of Prof Robert Jones, Chair, AHC FREC
CC: Professor Stephen Westland

Appendix H

CIELAB Values for 45 Samples on iiyama display

Stimuli Set	Image	L*	a*	b*
1	Colour Centre	81.0	-2.1	0.3
	2	82.8	-0.6	-1.2
	3	82.9	-1.0	2.2
	4	82.8	-3.6	-1.1
	5	82.9	-3.9	2.2
	6	79.4	-0.7	-1.4
	7	79.5	-1.0	2.0
	8	79.4	-3.6	-1.3
	9	79.5	-4.1	2.1
2	Colour Centre	74.3	-2.1	3.2
	2	75.9	0.0	1.0
	3	75.9	-0.7	4.8
	4	76.0	-3.3	1.4
	5	76.0	-3.5	4.9
	6	72.4	0.0	1.0
	7	72.4	-0.5	4.5
	8	72.7	-3.4	1.3
	9	72.6	-3.6	4.9
3	Colour Centre	70.3	-2.0	7.0
	2	72.0	-0.1	4.9
	3	72.0	-0.5	8.5
	4	72.2	-3.4	5.2
	5	72.2	-3.9	8.9
	6	68.5	-0.3	5.0
	7	68.5	-0.5	8.6
	8	68.5	-3.3	5.1
	9	68.6	-3.9	9.0
4	Colour Centre	68.8	-2.1	11.5
	2	70.6	-0.5	9.8
	3	70.5	-0.8	13.2
	4	70.7	-3.5	9.9
	5	70.7	-4.1	13.6
	6	67.0	-0.3	9.5
	7	67.0	-0.8	13.3
	8	67.1	-3.5	9.8
	9	67.0	-4.0	13.5
5	Colour Centre	65.3	-1.2	15.4
	2	67.1	0.7	13.3
	3	67.0	0.2	17.1
	4	67.3	-2.7	13.6
	5	67.2	-3.2	17.4
	6	63.5	0.7	13.4
	7	63.5	0.1	17.4
	8	63.6	-2.6	13.7
	9	63.6	-3.1	17.6

Appendix I

Participant Information Sheet

UNIVERSITY OF LEEDS

Participant Information Sheet

Measurement of Translucent Materials

Researcher: Chelsea Sullivan
Research Advisor: Professor Stephen Westland

You are being invited to take part in a research study. Before you decide, you need to understand what it would involve for you. Please take the time to read the following information. Please ask questions if anything you read is not clear or you would like more information. Take your time to decide whether or not you wish to participate.

The purpose of this project is to scale the perceptual yellowness of a set of teeth. The reason for doing this is to extract perceptual interval scale of yellowness values for each of the teeth that will be used to optimise a yellowness index. A yellowness index can help predict perceptual yellowness of teeth. This is important as to quantitatively evaluate the effectiveness of whitening toothpastes and other whitening treatments.

Do I have to take part?

It is completely up to you to decide. I will describe the study and walk you through the information sheet. If you decide to take park in the study, you will be given this information sheet to keep and will be asked to sign a consent form and right before the questionnaire. You are free to withdraw at any time without giving a reason.

What will happen to me if I take part?

If you take part in the study, it will take 40 minutes of your time and you will be reimbursed £10 for your time. The questionnaire will only happen once, and there will be no follow up. After discussing the research, you will be asked to sign a consent form. We will ask your age and gender. You will be asked to sit in front of a colour-calibrated computer in a darkened room to view a set of digital teeth samples. A set of teeth samples varying in colour will be displayed on the screen and you will be asked to rank order the images based on yellowness. To rank the images, you will use the mouse to drag images on the screen and place in a row them from left to right based on yellowness. Once the questionnaire is finished, you are free to leave.

Risk & Withdrawal:

There are no health risks associated with this research. All data you provide is strictly confidential. You will not be able to be identified in any reports or publications. If you withdraw from the study at any point during the questionnaire, all the information and data collected from you will be destroyed.

If you have any problems with this study or want further information here is contact details:

Researcher: Chelsea Sullivan Email: cp15crs@leeds.ac.uk

Appendix J

UNIVERSITY OF LEEDS

CONSENT FORM FOR PARTICPANT

School of Design PhD Project:

Measurement of Translucent Materials

Researcher: Chelsea Sullivan Research Advisor: Professor Stephen Westland

Please Initial the Following Boxes if you agree:

	I confirm that I have read and understand the information sheet about *Measurement of Translucent Materials* and discussed the project with Chelsea Sullivan who is conducting this research as part of a PhD in Design supervised by Professor Stephen Westland at the University of Leeds.
	I understand that my participation in this research is voluntary, I am free to refuse to participate and I am free to withdraw from the research at any time. My refusal to participate or withdrawal of consent will not affect my treatment in any way and there will be no negative consequences.
	I have been advised that there is no potential risks associated with this research. I give permission for members of the research team to have access to my anonymised responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report(s) that result from the research.
	I agree for the data collected from me to be stored and used in relevant future research in an anonymised form.
	I understand that other genuine researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.
	I understand that the data collected from my participation will be used for purpose (eg thesis, journal publication, reports, web pages, and other research outputs, etc) but only if they agree to preserve the confidentiality of the information as requested in this form. I consent for it to be used in that manner.
	I understand that relevant sections of the data collected during the study, may be looked at by individuals from the University of Leeds or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.

Name of participant:_____

Name of researcher: Chelsea Sullivan

Participant's signature:

_____

Researcher's signature:

_____

Appendix K

Calculated indices for 52 samples using the spectroradiometer (dif/0°)

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-8.32	21.70	23.69	26.21	30.71	-20.65	11.66
0M1	18.59	37.49	30.36	19.77	22.24	-36.88	7.57
0M2	9.83	32.28	27.47	22.15	25.62	-31.67	9.13
0M3	-5.75	23.01	24.21	25.55	29.89	-22.04	11.22
1.5M2	-50.97	-5.21	12.68	35.34	43.19	6.57	17.15
1M1	-20.13	14.42	20.47	28.92	34.43	-13.29	13.37
1M1.5	-39.97	2.83	15.24	33.64	40.84	-1.51	16.47
1M2	-41.42	2.56	15.90	33.83	40.69	-0.95	16.54
2.5M2	-56.87	-11.01	9.42	36.37	45.42	11.82	17.60
2L1.5	-44.69	-0.70	14.85	34.04	41.16	2.19	16.41
2L2.5	-71.12	-14.74	8.90	40.35	49.37	16.77	20.76
2M1	-29.61	6.35	17.81	30.01	36.14	-5.34	13.61
2M2	-63.54	-11.07	9.68	38.78	47.68	12.71	19.76
2M3	-66.71	-15.19	7.44	39.10	48.83	16.33	19.69
2R1.5	-37.26	2.18	14.61	32.50	39.92	-1.36	15.47
2R2.5	-43.97	-1.53	13.59	33.88	41.55	2.59	16.29
3.5M2	-67.23	-21.56	5.06	37.55	48.12	21.63	17.52
3L1.5	-43.24	-3.21	14.24	32.51	39.67	4.22	14.82
3L2.5	-76.71	-21.85	6.46	40.41	50.19	23.34	19.90
3M1	-47.44	-6.69	12.63	33.26	40.96	7.53	15.14
3M2	-68.60	-18.56	6.29	38.85	48.94	19.35	19.05
3M3	-72.45	-20.79	5.41	39.70	50.08	21.63	19.58
3R1.5	-53.53	-13.13	7.63	34.86	44.66	12.96	16.16
3R12.5	-74.98	-25.35	1.90	40.16	52.04	25.24	19.72
4.5M2	-73.41	-27.95	2.79	38.14	49.49	27.62	17.23
4L2.5	-79.89	-30.50	-0.37	40.81	53.57	29.95	19.73
4L21.5	-70.45	-24.71	4.75	37.55	48.06	24.81	16.99
4M1	-46.93	-13.02	8.57	31.81	41.07	12.35	13.52
4M2	-82.79	-32.77	0.29	40.70	53.04	32.49	19.07
4M3	-83.87	-33.21	-1.25	41.54	54.60	32.65	20.05
4R1.5	-59.80	-22.16	3.48	35.07	46.46	20.94	15.47
4R2.5	-79.67	-33.40	-1.39	40.00	53.23	32.27	18.59
5M1	-56.22	-19.55	7.57	32.92	42.18	19.20	13.51
5M2	-80.81	-34.66	-0.79	39.66	52.55	33.69	17.92
5M2.5	-82.39	-35.39	-1.58	40.29	53.53	34.34	18.49
5M3	-128.97	-62.21	-12.98	51.07	68.56	61.24	25.64
A1	-39.21	4.45	16.41	33.62	40.38	-2.83	16.55
A2	-60.06	-13.92	6.96	37.32	47.38	14.28	18.31
A3	-68.45	-18.22	5.67	39.18	49.59	18.89	19.50
A3.5	-78.10	-26.29	2.47	40.70	52.23	26.60	19.98
A4	-73.90	-27.05	2.99	38.62	49.91	26.91	17.81
B1	-29.44	9.64	20.63	30.38	35.44	-7.77	14.04
B2	-57.09	-8.06	11.91	36.76	44.79	9.71	18.09
B3	-83.39	-26.02	3.63	42.26	53.16	27.26	21.36
B4	-98.01	-35.81	-1.21	45.48	58.14	36.73	23.45
C1	-42.06	0.68	17.74	32.39	38.12	1.32	14.87
C2	-56.08	-11.59	10.61	35.30	43.70	12.55	16.41
C3	-65.93	-17.56	9.47	36.92	45.45	18.90	17.00
C4	-74.41	-25.77	6.04	38.01	47.75	26.57	16.99
D2	-43.81	-2.36	15.77	32.49	38.94	3.86	14.75
D3	-60.44	-16.34	7.88	36.02	45.48	16.75	16.64
D4	-66.12	-14.22	10.56	38.03	46.29	16.10	18.44

Appendix L

Calculated indices for 52 samples with the spectroradiometer (45°/0°)

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-9.35	19.75	23.13	25.88	30.44	-18.83	11.27
0M1	17.05	35.17	29.55	19.53	22.10	-34.70	7.30
0M2	7.12	28.98	26.55	21.98	25.54	-28.49	8.84
0M3	-5.99	21.62	23.72	25.14	29.54	-20.82	10.82
1.5M2	-51.83	-6.99	12.41	34.97	42.82	8.25	16.60
1M1	-21.47	12.35	19.99	28.65	34.18	-11.32	12.97
1M1.5	-43.10	-0.24	14.42	33.79	41.15	1.49	16.29
1M2	-44.45	-0.60	15.14	33.87	40.86	2.14	16.24
2.5M2	-60.75	-14.62	8.10	36.82	46.29	15.27	17.60
2L1.5	-49.02	-4.72	13.41	34.50	42.01	6.06	16.40
2L2.5	-72.94	-16.80	8.48	40.33	49.43	18.77	20.44
2M1	-31.10	4.12	17.13	29.81	36.06	-3.25	13.25
2M2	-64.44	-12.65	9.35	38.58	47.54	14.18	19.36
2M3	-70.08	-18.23	6.54	39.44	49.41	19.28	19.62
2R1.5	-42.96	-2.79	13.15	33.07	40.83	3.54	15.47
2R2.5	-44.14	-2.94	13.09	33.51	41.30	3.79	15.83
3.5M2	-69.74	-23.93	4.17	37.87	48.75	23.87	17.52
3L1.5	-51.15	-9.43	11.99	33.73	41.54	10.34	15.21
3L2.5	-80.63	-25.12	5.24	40.99	51.18	26.49	20.03
3M1	-48.37	-8.15	11.85	33.28	41.26	8.80	15.04
3M2	-70.45	-20.94	5.58	38.82	49.13	21.57	18.71
3M3	-77.76	-25.08	3.90	40.50	51.38	25.80	19.78
3R1.5	-56.47	-15.89	6.62	35.20	45.34	15.58	16.17
3R12.5	-82.97	-30.88	-0.16	41.69	54.28	30.71	20.47
4.5M2	-77.28	-30.94	1.60	38.85	50.67	30.50	17.50
4L2.5	-86.28	-35.47	-2.26	41.94	55.41	34.73	20.16
4L21.5	-73.80	-27.39	3.60	38.17	49.13	27.36	17.23
4M1	-48.68	-14.97	7.56	32.05	41.72	14.10	13.55
4M2	-87.93	-36.48	-1.35	41.77	54.76	36.05	19.63
4M3	-89.74	-37.84	-3.01	42.56	56.30	37.10	20.41
4R1.5	-62.80	-24.98	2.27	35.54	47.40	23.56	15.58
4R2.5	-82.00	-35.78	-2.44	40.34	54.01	34.44	18.62
5M1	-62.55	-24.30	5.50	34.24	44.32	23.75	14.07
5M2	-84.95	-38.08	-2.25	40.45	53.94	36.90	18.20
5M2.5	-86.47	-38.84	-2.94	40.99	54.79	37.60	18.69
5M3	-133.85	-65.82	-14.63	52.09	70.27	64.64	26.22
A1	-40.03	2.54	15.57	33.40	40.41	-1.16	16.19
A2	-61.83	-16.13	6.04	37.42	47.80	16.29	18.17
A3	-71.91	-21.60	4.40	39.57	50.41	22.06	19.46
A3.5	-80.56	-28.92	1.54	40.92	52.78	29.05	19.83
A4	-77.71	-30.25	1.61	39.30	51.13	29.93	18.04
B1	-31.20	7.31	19.47	30.45	35.86	-5.67	13.93
B2	-59.04	-10.42	10.97	36.84	45.16	11.89	17.92
B3	-85.71	-28.50	2.93	42.39	53.51	29.61	21.11
B4	-99.35	-37.41	-1.71	45.54	58.39	38.21	23.26
C1	-45.31	-2.52	16.28	32.79	38.97	4.31	14.92
C2	-63.13	-17.04	8.43	36.52	45.64	17.84	16.89
C3	-70.44	-21.22	7.84	37.73	46.85	22.38	17.31
C4	-79.66	-29.66	4.27	39.10	49.53	30.29	17.50
D2	-46.47	-5.10	14.47	32.81	39.68	6.40	14.78
D3	-62.31	-18.33	6.92	36.27	46.09	18.56	16.67
D4	-69.70	-17.39	9.17	38.58	47.31	19.09	18.59

Appendix M

Calculated indices for 52 samples using the spectrophotometer (SPI)

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-4.43	18.32	23.98	22.46	26.14	-17.77	8.56
0M1	16.73	30.59	29.01	17.20	19.25	-30.41	5.50
0M2	5.68	22.89	25.91	19.41	22.32	-22.67	6.67
0M3	-3.18	18.84	24.00	22.17	25.87	-18.40	8.40
1.5M2	-34.73	-1.58	16.37	28.71	34.64	2.35	11.75
1M1	-7.64	17.62	23.44	23.80	27.81	-16.92	9.51
1M1.5	-29.19	3.54	17.68	28.43	34.24	-2.76	12.04
1M2	-31.09	2.17	17.89	28.42	33.95	-1.21	11.84
2.5M2	-39.01	-5.71	13.73	29.68	36.69	6.00	12.24
2L1.5	-34.43	-1.69	16.62	28.39	34.13	2.50	11.46
2L2.5	-51.96	-8.57	13.58	33.75	40.80	10.03	15.16
2M1	-18.78	7.00	19.80	24.61	29.31	-6.55	9.37
2M2	-46.17	-5.35	14.24	32.67	39.63	6.54	14.65
2M3	-50.24	-10.51	11.61	33.00	40.88	11.15	14.48
2R1.5	-27.20	2.40	17.12	27.19	33.07	-2.02	11.01
2R2.5	-30.70	1.35	16.72	28.37	34.44	-0.77	11.82
3.5M2	-44.74	-10.86	11.45	30.56	38.36	10.87	12.47
3L1.5	-30.42	-0.73	16.97	26.92	32.43	1.29	10.44
3L2.5	-56.68	-13.60	11.54	34.14	41.81	14.73	14.95
3M1	-31.36	-0.48	16.47	27.77	33.70	0.98	11.16
3M2	-50.93	-11.73	11.23	32.85	40.82	12.27	14.22
3M3	-52.70	-13.94	10.50	32.83	40.97	14.35	13.95
3R1.5	-35.98	-5.46	12.82	28.94	36.44	5.21	11.81
3R12.5	-56.29	-16.79	8.01	34.10	43.42	16.74	14.89
4.5M2	-51.02	-16.32	9.14	31.57	40.17	16.05	12.73
4L2.5	-61.80	-22.89	5.30	34.71	45.05	22.30	14.77
4L21.5	-49.13	-14.20	10.33	31.26	39.37	14.18	12.65
4M1	-30.79	-5.80	12.88	26.38	33.58	5.04	9.84
4M2	-63.36	-21.98	6.11	35.38	45.34	21.85	15.37
4M3	-63.51	-24.03	5.10	34.96	45.31	23.50	14.80
4R1.5	-39.72	-12.83	9.34	28.39	37.04	11.64	10.79
4R2.5	-57.96	-21.94	5.43	33.53	43.84	21.03	13.93
5M1	-41.41	-12.24	11.36	28.34	35.79	11.86	10.56
5M2	-57.33	-22.13	6.00	32.89	42.83	21.30	13.27
5M2.5	-58.82	-23.51	5.41	33.12	43.29	22.59	13.31
5M3	-99.77	-44.66	-4.22	44.04	58.11	44.11	20.60
A1	-25.50	6.26	19.24	27.53	32.74	-5.32	11.51
A2	-40.67	-6.94	12.45	30.37	38.01	7.00	12.77
A3	-48.02	-10.70	11.30	32.12	40.12	11.02	13.77
A3.5	-53.32	-15.47	9.12	32.99	41.77	15.48	14.03
A4	-44.66	-12.99	10.57	29.84	37.84	12.64	11.70
B1	-18.04	10.41	22.01	25.13	29.12	-9.32	9.92
B2	-38.64	-2.02	16.16	30.29	36.42	3.11	12.95
B3	-59.75	-15.66	10.19	35.02	43.27	16.64	15.55
B4	-70.76	-22.30	6.65	37.89	47.56	23.12	17.44
C1	-27.58	3.81	19.90	26.73	31.16	-2.57	10.55
C2	-38.21	-5.42	14.73	28.95	35.36	5.93	11.60
C3	-45.06	-8.94	13.97	30.43	36.96	9.78	12.32
C4	-50.84	-14.26	11.53	31.31	38.70	14.77	12.51
D2	-26.92	1.93	18.72	25.93	30.67	-1.13	9.85
D3	-39.47	-7.14	13.38	29.27	36.28	7.32	11.77
D4	-47.04	-7.02	14.85	31.87	38.27	8.39	13.67

Appendix N

Calculated indices for 52 samples using the spectrophotometer (SPE)

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-14.58	11.08	21.57	24.15	28.36	-10.43	9.31
0M1	9.74	25.13	27.22	18.13	20.48	-24.89	5.90
0M2	-3.40	16.16	23.62	20.83	24.21	-15.86	7.28
0M3	-13.03	11.77	21.65	23.80	28.02	-11.23	9.11
1.5M2	-50.54	-11.89	12.85	31.78	38.72	12.84	13.09
1M1	-17.95	10.31	21.07	25.51	30.04	-9.49	10.26
1M1.5	-42.59	-5.43	14.67	30.91	37.52	6.36	13.14
1M2	-44.91	-6.93	14.84	31.01	37.36	8.08	12.99
2.5M2	-55.62	-16.67	9.92	32.94	41.13	17.08	13.66
2L1.5	-49.75	-11.66	13.21	31.36	38.08	12.65	12.74
2L2.5	-67.62	-18.50	10.32	36.74	44.72	20.19	16.53
2M1	-31.39	-1.58	16.83	26.94	32.42	2.16	10.36
2M2	-62.25	-15.68	10.79	35.73	43.68	17.09	16.08
2M3	-68.26	-22.12	7.66	36.53	45.64	22.93	16.10
2R1.5	-43.07	-8.23	13.43	30.23	37.16	8.74	12.35
2R2.5	-41.35	-6.02	14.29	30.24	36.95	6.69	12.58
3.5M2	-62.88	-22.72	7.27	34.25	43.43	22.81	14.07
3L1.5	-43.93	-9.65	13.87	29.56	35.95	10.34	11.54
3L2.5	-75.13	-25.17	7.65	37.81	46.68	26.53	16.62
3M1	-45.62	-9.94	13.21	30.51	37.37	10.59	12.34
3M2	-67.60	-22.48	7.56	36.15	45.27	23.17	15.70
3M3	-69.62	-24.84	6.74	36.22	45.59	25.37	15.45
3R1.5	-50.32	-15.17	9.37	31.75	40.36	14.95	13.03
3R12.5	-72.84	-27.70	4.21	37.38	47.98	27.69	16.36
4.5M2	-70.61	-29.12	4.60	35.63	45.84	28.88	14.45
4L2.5	-81.32	-35.67	0.79	38.73	50.72	35.06	16.54
4L21.5	-67.50	-26.10	6.15	35.03	44.58	26.17	14.27
4M1	-45.03	-15.68	9.24	29.23	37.65	14.87	10.99
4M2	-82.41	-34.27	1.83	39.27	50.74	34.18	17.13
4M3	-83.61	-37.15	0.47	39.11	51.19	36.60	16.62
4R1.5	-58.02	-25.44	4.68	32.19	42.57	24.11	12.32
4R2.5	-79.16	-35.98	0.41	37.95	50.16	34.99	15.85
5M1	-58.60	-23.68	7.23	31.92	40.82	23.29	11.96
5M2	-79.00	-36.39	0.90	37.46	49.34	35.49	15.21
5M2.5	-81.41	-38.38	0.09	37.91	50.15	37.37	15.33
5M3	-128.73	-62.87	-10.53	49.94	66.43	62.30	23.54
A1	-38.90	-2.71	16.25	29.99	35.96	3.82	12.61
A2	-56.12	-17.27	8.87	33.38	42.15	17.40	14.09
A3	-64.08	-21.20	7.67	35.29	44.44	21.62	15.18
A3.5	-71.98	-27.59	4.90	36.76	46.97	27.69	15.70
A4	-64.03	-25.78	6.00	33.83	43.43	25.46	13.35
B1	-30.93	1.77	19.13	27.47	32.14	-0.48	10.93
B2	-54.65	-12.35	12.70	33.36	40.46	13.66	14.34
B3	-79.19	-27.85	6.08	38.89	48.44	29.05	17.34
B4	-92.63	-35.97	2.02	42.25	53.44	36.98	19.54
C1	-41.39	-5.18	16.89	29.34	34.53	6.65	11.66
C2	-53.77	-15.56	11.22	32.04	39.52	16.22	12.92
C3	-62.03	-19.72	10.27	33.84	41.51	20.76	13.78
C4	-69.15	-25.92	7.51	35.07	43.78	26.58	14.06
D2	-41.10	-7.37	15.51	28.67	34.28	8.36	10.99
D3	-53.06	-16.23	10.21	31.93	39.92	16.48	12.87
D4	-62.98	-17.07	11.52	34.98	42.33	18.68	15.06

Appendix O

Calculated indices for 52 samples using spectrophotometer (0°/45°)

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-8.17	14.69	24.12	21.95	24.97	-13.84	7.90
0M1	13.62	26.93	29.02	16.37	17.74	-26.51	4.79
0M2	4.67	20.77	26.42	18.26	20.38	-20.34	5.75
0M3	-7.67	14.92	23.89	21.98	25.17	-14.18	7.95
1.5M2	-41.91	-6.43	16.08	29.28	34.77	7.68	11.53
1M1	-13.26	13.36	23.15	24.08	27.72	-12.33	9.35
1M1.5	-36.27	-1.20	17.44	29.05	34.46	2.44	11.95
1M2	-38.49	-2.82	17.58	29.04	34.14	4.27	11.69
2.5M2	-49.59	-12.37	12.63	31.22	38.17	13.15	12.58
2L1.5	-41.87	-6.77	16.18	28.99	34.32	8.05	11.24
2L2.5	-62.56	-14.87	12.85	35.26	42.12	16.92	15.49
2M1	-25.97	1.63	19.12	25.04	29.34	-0.78	9.13
2M2	-56.73	-11.73	13.37	34.21	41.07	13.47	15.07
2M3	-61.88	-17.97	10.48	34.60	42.43	19.12	14.73
2R1.5	-35.15	-3.28	16.50	27.84	33.37	4.08	10.84
2R2.5	-38.27	-3.82	16.15	29.14	34.95	4.80	11.80
3.5M2	-56.57	-18.33	9.79	32.53	40.56	18.73	13.04
3L1.5	-39.06	-6.49	15.85	28.06	33.51	7.41	10.61
3L2.5	-70.54	-21.67	9.87	36.57	44.45	23.37	15.82
3M1	-40.47	-6.29	15.36	29.12	35.05	7.19	11.52
3M2	-63.06	-19.08	9.70	34.92	43.09	20.08	14.92
3M3	-64.34	-20.85	9.36	34.71	42.87	21.80	14.46
3R1.5	-48.41	-13.34	10.93	31.09	38.94	13.46	12.58
3R12.5	-71.61	-25.59	5.90	37.07	46.88	26.06	16.19
4.5M2	-65.39	-25.31	6.79	34.21	43.39	25.39	13.61
4L2.5	-63.33	-22.82	8.07	33.95	42.61	23.19	13.66
4L21.5	-77.54	-32.47	3.01	37.62	48.55	32.31	15.79
4M1	-40.88	-12.54	11.07	28.09	35.65	12.01	10.36
4M2	-78.85	-31.74	3.47	38.24	48.96	31.90	16.42
4M3	-76.75	-32.20	3.37	37.25	47.97	32.07	15.44
4R1.5	-51.84	-20.85	7.24	30.51	39.66	19.90	11.40
4R2.5	-73.20	-31.77	2.81	36.33	47.41	31.12	14.84
5M1	-55.89	-21.48	8.53	31.17	39.43	21.33	11.55
5M2	-76.48	-33.81	2.57	36.80	47.87	33.30	14.85
5M2.5	-76.45	-34.41	2.39	36.59	47.71	33.81	14.57
5M3	-126.55	-60.38	-8.90	49.41	65.12	60.24	23.23
A1	-31.41	1.86	19.25	27.68	32.31	-0.49	11.09
A2	-50.87	-13.76	11.37	31.69	39.27	14.24	12.92
A3	-57.51	-16.86	10.51	33.33	41.17	17.63	13.85
A3.5	-67.80	-24.07	7.28	35.59	44.69	24.61	14.95
A4	-59.30	-22.01	8.27	32.54	41.05	22.08	12.61
B1	-23.62	6.32	21.83	25.34	28.84	-4.83	9.64
B2	-49.39	-8.73	15.07	31.84	37.89	10.33	13.34
B3	-74.37	-24.21	8.47	37.56	46.04	25.78	16.45
B4	-86.93	-31.87	4.66	40.69	50.73	33.26	18.43
C1	-35.36	-1.21	19.21	27.64	31.80	2.90	10.64
C2	-49.17	-12.28	13.31	30.69	37.17	13.24	12.07
C3	-57.39	-16.40	12.24	32.56	39.32	17.71	13.01
C4	-66.28	-23.58	8.90	34.31	42.37	24.49	13.63
D2	-35.16	-3.52	17.85	26.87	31.35	4.77	9.90
D3	-49.28	-13.53	12.10	30.71	37.78	14.09	12.06
D4	-58.18	-14.02	13.36	33.61	40.14	15.81	14.15

Appendix P

Calculated indices for 52 samples using the Vita EasyShade

SHADE	WIO	WIC	WID	YIE313	YID1925	YIO	b*
0.5M1	13.14	41.39	26.94	25.54	30.11	-40.58	12.39
0M1	35.91	55.71	32.77	20.62	23.63	-55.22	8.99
0M2	20.81	43.25	28.85	22.43	26.23	-42.79	9.92
0M3	15.23	43.00	26.77	25.48	30.22	-42.33	12.45
1.5M2	-21.63	18.41	18.93	32.03	38.60	-17.13	16.50
1M1	0.19	34.86	23.13	29.21	34.99	-33.84	15.21
1M1.5	-19.27	21.35	18.13	32.68	39.76	-20.23	17.41
1M2	-20.67	19.67	19.08	32.15	38.74	-18.35	16.72
2.5M2	-35.79	7.01	12.84	34.76	43.30	-6.32	18.18
2L1.5	-29.24	10.69	17.28	32.17	38.96	-9.52	15.94
2L2.5	-69.86	-14.17	8.65	40.22	49.40	16.04	20.77
2M1	-5.74	25.84	22.21	27.76	33.33	-25.09	13.33
2M2	-41.30	9.88	12.79	38.06	46.68	-8.26	21.61
2M3	-39.99	6.89	11.78	36.69	45.64	-5.92	20.03
2R1.5	-16.86	19.95	18.41	30.99	37.88	-19.19	15.78
2R2.5	-22.09	18.04	16.85	32.78	40.22	-17.14	17.27
3.5M2	-59.56	-10.79	5.59	38.94	49.74	11.10	20.44
3L1.5	-35.24	5.52	15.01	33.07	40.56	-4.57	16.33
3L2.5	-63.34	-6.90	7.04	41.34	51.53	8.38	23.18
3M1	-40.18	4.18	12.23	35.46	44.11	-3.36	18.53
3M2	-60.25	-7.57	6.02	40.32	50.98	8.39	22.18
3M3	-58.00	-7.23	7.41	39.15	49.28	8.09	20.96
3R1.5	-57.65	-13.28	5.93	37.08	47.68	13.19	18.31
3R2.5	-82.50	-26.98	-0.21	43.01	55.65	27.17	22.50
4.5M2	-67.88	-18.33	2.22	40.22	52.21	18.17	20.91
4L1.5	-76.36	-22.97	-1.16	42.68	55.88	22.65	22.99
4L2.5	-62.36	-13.26	4.44	39.40	50.60	13.42	20.64
4M1	-38.58	-2.01	8.86	33.39	43.27	1.36	16.26
4M2	-95.35	-33.23	-7.34	47.57	62.80	32.88	27.14
4M3	-77.64	-23.45	-1.17	42.96	56.13	23.23	23.20
4R1.5	-52.60	-12.59	4.26	36.07	47.52	11.61	17.70
4R2.5	-75.64	-24.74	-2.38	42.20	55.95	23.88	22.38
5M1	-64.98	-19.51	3.31	38.27	49.79	19.11	18.76
5M2	-81.02	-28.71	-3.61	43.07	57.23	27.81	22.74
5M2.5	-75.78	-24.70	-2.36	42.26	56.00	23.87	22.44
5M3	-109.99	-43.10	-14.44	51.29	68.99	42.04	30.50
A1	-38.60	8.59	10.01	37.45	47.19	-7.95	21.10
A2	-24.39	17.32	14.88	34.05	42.26	-16.60	18.48
A3	-39.72	6.75	10.18	37.07	46.68	-6.14	20.51
A3.5	-56.44	-5.89	6.11	39.56	50.30	6.44	21.67
A4	-61.69	-13.18	4.81	39.05	50.10	13.34	20.28
B1	-22.66	15.70	18.15	31.49	38.22	-14.68	15.81
B2	-27.36	16.08	17.36	33.82	40.93	-14.64	17.97
B3	-57.56	-2.86	8.54	40.34	50.10	4.34	22.60
B4	-68.06	-8.70	4.02	43.31	54.59	9.98	25.26
C1	-41.03	7.06	13.77	36.47	44.59	-5.56	19.67
C2	-32.99	9.25	15.60	33.64	41.11	-8.12	17.20
C3	-45.01	0.93	11.94	36.03	44.62	0.10	18.67
C4	-62.08	-11.39	6.96	39.13	49.28	12.23	20.38
D2	-55.20	-7.63	9.18	37.29	46.65	8.49	18.96
D3	-31.83	7.66	15.86	32.35	39.59	-6.74	15.88
D4	-46.78	0.40	10.61	36.96	46.10	0.52	19.59

Appendix Q

Calculated indices for 52 samples using the DigiEye + NikonD7000

SHADE	WIC	WIO	WID	YIE313	YID1925	YIO	b*
0.5M1	-6.74	24.02	22.03	27.36	32.86	-23.33	12.87
0M1	19.87	40.88	29.64	21.32	24.53	-40.34	8.91
0M2	9.78	34.76	26.52	23.86	28.07	-34.18	10.63
0M3	-4.07	25.53	22.41	26.82	32.24	-24.95	12.53
1.5M2	-51.74	-6.34	10.60	35.97	44.76	7.22	17.79
1M1	-17.08	18.09	19.76	29.71	35.86	-17.18	14.42
1M1.5	-40.80	3.36	12.92	35.09	43.40	-2.40	18.04
1M2	-40.73	4.45	14.77	34.85	42.38	-3.01	17.79
2.5M2	-57.57	-11.08	7.45	37.39	47.37	11.52	18.74
2L1.5	-45.26	-0.32	13.79	34.87	42.54	1.67	17.29
2L2.5	-68.78	-12.97	8.03	40.55	50.07	14.69	21.36
2M1	-23.75	10.90	17.87	29.78	36.25	-10.20	13.90
2M2	-64.05	-10.63	8.51	39.60	49.05	12.11	20.76
2M3	-64.72	-13.14	7.20	39.31	49.25	14.18	20.24
2R1.5	-35.91	4.38	13.17	33.53	41.79	-3.85	16.76
2R2.5	-43.02	0.53	12.49	34.88	43.21	0.35	17.53
3.5M2	-65.63	-19.77	4.24	38.04	49.10	19.63	18.38
3L1.5	-47.66	-4.86	12.47	34.28	42.24	5.82	16.29
3L2.5	-78.00	-20.84	5.47	41.68	51.96	22.35	21.47
3M1	-41.64	-2.40	12.90	32.83	40.74	2.97	15.34
3M2	-63.93	-15.42	5.23	38.82	49.59	15.72	19.67
3M3	-72.72	-21.03	4.07	40.23	51.24	21.58	20.23
3R1.5	-49.05	-10.29	6.55	34.79	45.29	9.60	16.64
3R12.5	-75.96	-25.76	0.89	40.75	53.06	25.51	20.35
4.5M2	-65.09	-21.99	3.53	37.19	48.52	21.41	17.34
4L2.5	-76.79	-28.39	-2.06	41.14	54.76	27.31	20.69
4L21.5	-71.78	-25.45	3.65	38.21	49.27	25.37	17.63
4M1	-46.36	-12.96	7.44	32.12	42.00	11.95	13.91
4M2	-76.84	-27.90	0.50	40.38	52.83	27.44	19.68
4M3	-82.37	-31.82	-1.76	41.73	55.07	31.12	20.59
4R1.5	-59.60	-22.65	1.93	35.45	47.61	20.98	15.93
4R2.5	-74.70	-29.99	-2.00	39.74	53.40	28.49	19.08
5M1	-63.92	-23.31	5.30	35.36	45.67	22.93	15.24
5M2	-77.63	-32.80	-1.65	39.59	53.00	31.43	18.37
5M2.5	-83.30	-36.66	-3.21	40.79	54.80	35.20	19.00
5M3	-129.19	-62.75	-15.24	51.74	70.18	61.27	26.78
A1	-38.57	5.30	15.01	34.30	41.78	-3.99	17.37
A2	-58.21	-13.13	5.59	37.49	48.24	13.04	18.74
A3	-63.94	-14.62	4.42	39.45	50.60	14.82	20.48
A3.5	-78.03	-26.70	0.78	41.17	53.48	26.57	20.61
A4	-70.53	-24.51	2.24	38.72	50.53	24.03	18.48
B1	-27.58	12.44	19.38	31.50	37.44	-10.86	15.39
B2	-54.10	-4.93	11.95	36.98	45.21	6.50	18.75
B3	-84.99	-25.51	2.98	43.27	54.45	26.85	22.56
B4	-99.83	-35.48	-1.76	46.41	59.26	36.56	24.59
C1	-41.89	1.97	17.07	33.26	39.49	-0.08	15.85
C2	-53.85	-8.98	10.85	35.52	43.99	9.94	16.97
C3	-64.39	-15.59	9.52	37.19	45.87	16.92	17.58
C4	-73.72	-24.69	5.93	38.29	48.17	25.47	17.47
D2	-42.83	-0.95	15.32	32.97	39.81	2.33	15.40
D3	-53.96	-11.74	8.05	35.47	45.15	11.83	16.86
D4	-68.77	-13.71	10.07	39.46	47.99	15.82	19.92

Appendix R

Table A: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the spectroradiometer (diffuse/0°) to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1146	1326	13.57	161	1326	12.14
WIO	1228	1326	7.39	77	1326	5.81
WID	1203	1326	9.28	95	1326	7.16
YIE313	275	1326	20.74	1079	1326	18.63
YID1925	210	1326	15.84	1146	1326	13.57
YIO	108	1326	8.14	1235	1326	6.86
b*	375	1326	28.28	976	1326	26.40

Table B: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the spectroradiometer (45°/0°) to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1156	1326	12.82051	155	1326	11.69
WIO	1236	1326	6.78733	70	1326	5.28
WID	1216	1326	8.295626	84	1326	6.33
YIE313	247	1326	18.62745	1100	1326	17.04
YID1925	191	1326	14.40422	1162	1326	12.37
YIO	101	1326	7.616893	1239	1326	6.56
b*	341	1326	25.71644	1011	1326	23.76

Table C: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the spectrophotometer with the specular component included compared to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1121	1326	15.46	1145	1326	13.65
WIO	1211	1326	8.67	1234	1326	6.94
WID	1203	1326	9.28	1229	1326	7.32
YIE313	282	1326	21.27	259	1326	19.53
YID1925	224	1326	16.89	200	1326	15.08
YIO	132	1326	9.95	113	1326	8.52
b*	357	1326	26.92	333	1326	25.11

Table D: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the spectrophotometer with the specular component excluded compared to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1135	1326	14.40	1157	1326	12.75
WIO	1214	1326	8.45	1236	1326	6.79
WID	1216	1326	8.30	1243	1326	6.26
YIE313	247	1326	18.63	225	1326	16.97
YID1925	200	1326	15.08	177	1326	13.35
YIO	125	1326	9.43	108	1326	8.14
b*	327	1326	24.66	306	1326	23.08

Table E: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the spectrophotometer (0°/45°) compared to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1148	1326	13.42	161	1326	12.14
WIO	1227	1326	7.47	78	1326	5.88
WID	1238	1326	6.64	63	1326	4.75
YIE313	230	1326	17.35	1115	1326	15.91
YID1925	181	1326	13.65	1166	1326	12.07
YIO	113	1326	8.52	1232	1326	7.09
b*	299	1326	22.55	1046	1326	21.12

Table F: The % Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the Vita EasyShade compared to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1142	1326	13.88	1141	1326	13.95
WIO	1168	1326	11.92	1162	1326	12.37
WID	1180	1326	11.01	1178	1326	11.16
YIE313	229	1326	17.27	216	1326	16.29
YID1925	193	1326	14.56	180	1326	13.57
YIO	158	1326	11.92	165	1326	12.44
b*	284	1326	21.42	275	1326	20.74

Table G: Wrong Decision for the various indices calculated from measurements collected of the 52 samples from the DigiEye + NikonD7000 with diffuse/0° geometry compared to the visual perception of samples based on whiteness as well as yellowness

	Whiteness			Yellowness		
Indices	Correct	Total Comparisons	%WD	Correct	Total Comparisons	%WD
WIC	1135	1326	14.40	1157	1326	12.75
WIO	1230	1326	7.24	1242	1326	6.33
WID	1190	1326	10.26	1219	1326	8.07
YIE313	292	1326	22.02	267	1326	20.14
YID1925	222	1326	16.74	192	1326	14.48
YIO	111	1326	8.37	99	1326	7.47
b*	410	1326	30.92	383	1326	28.88