# Human Visual Perception of Haze in White Wine and Model Solutions and Relationships with Instrumental Turbidity and Imaging Models

---

A thesis

submitted in partial fulfilment

of the requirements for the Degree of

Master of Applied Science

at

Lincoln University

by

Joseph Falconer

---

Lincoln University

2014

Abstract of a thesis submitted in partial fulfilment of the
requirements for the Degree of Master of Applied Science.

# Human Visual Perception of Haze in White Wine and Relationships with Instrumental Turbidity and Imaging Models

by

Joseph Falconer

White wines are liable to develop haze during storage and transport due to residual protein in the finished wine. Since haze in white wines is considered unacceptable, winemakers attempt to reduce the protein content prior to bottling through the use of fining agents such as bentonite. The use of bentonite is not without economic cost with the estimated annual impact to the New Zealand wine industry in the region of NZD 35 million. The optimum quantity of bentonite to be used is normally determined using a heat test in conjunction with empirical fining trials. During these tests the winemaker is required to conduct various haze comparisons. Since these assessments are difficult to perform visually, researchers recommend the use of suitable nephelometric instrumentation to ensure consistency and precision. Despite these recommendations, most winemakers have not purchased the necessary equipment; it is thought that cost is the primary deterrent.

The overall goal of this project was to explore the relationships between nephelometry, digital imaging and human perception of haze in terms of thresholds and intensity. In addition, the feasibility of using digital imaging devices to make nephelometric measurements was to be investigated.

A range of achromatic and yellow turbid suspensions were prepared by suspending synthetic polymer microspheres of diameter 0.25 µm in aqueous solutions. Base wine stocks exhibiting high turbidity were prepared by heating commercial grade Chardonnay wine at 90˚ for two hours. A range of wine samples of varying turbidities were then prepared by dilution with untreated wine. Polymer microsphere suspensions and wine samples of varying visual haze were used in a series of sensory and instrumentation experiments. The sensory experiments examined haze detection thresholds and intensity scaling in human subjects. All evaluations took place in purpose designed sensory booths utilising overhead light emitting diode (LED) illumination. The booth backgrounds and sides were

modified to present dark or light surrounds by hanging black or white linen on the walls. Samples were presented to human subjects in standard ISO tasting glasses and the subjects were permitted to manipulate the glass freely whilst they evaluated the samples. Turbidity measurements were made with a Hach 2100P nephelometric device. Image data was captured using a modified Canon A2300 digital still camera using 90˚ viewing geometries.

Haze threshold determinations were made for achromatic and yellow microsphere suspensions using the Ascending Method of Limits (AML) and 3-Alternate Forced Choice (3AFC) methods. The lowest thresholds were found when achromatic suspensions were presented in a dark surround where individual thresholds ranged from 0.21 Nephelometric Turbidity Units (NTU) to 1.07 NTU. The highest thresholds were found when yellow suspensions were presented in a light surround where individual thresholds ranged from 1.89 NTU to 25.16 NTU. Heat treated Chardonnay wine samples were also evaluated in a dark surround where individual thresholds ranged from 0.52 NTU to 1.39 NTU.

Visual haze intensities in achromatic and yellow microsphere suspensions were assessed by human subjects in dark and light surrounds using modulus Magnitude Estimation (ME). Sample luminance was determined by digitally photographing the samples under the same conditions employed for the ME evaluations. The perceived haze intensity with respect to luminance was found to follow generalised power or logarithmic functions similar to psycho-physical models commonly proposed for the response of the human visual system to brightness and lightness. The data exhibited a bi-segment nature indicative of surround and/or planar contrast induction. Similar responses were found for heat treated Chardonnay wine samples using turbidity as the independent variable.

Turbidity values were measured for a range achromatic and yellow microsphere suspension samples. Luminance values were derived from digital images taken of ISO tasting glasses containing the same microsphere suspensions. The turbidity and luminance values were found to be linearly related below 50 NTU (dark surround, $R^2$=0. 9978; light surround, $R^2$=0. 9813). A subsequent experiment examined turbidity values and luminance data in achromatic and yellow microsphere suspension samples over a low range (< 6 NTU) by imaging the sample surface directly rather than through the glass receptacle. Turbidity vales and luminance data were again found to be linearly related ($R^2$=0.9628). Measurements from the Hach 2100P device and a luminance based measurement model were found to be mean and median equivalent for the low range experimental data.

**Keywords:** Protein stabilisation; Heat test; Nephelometer; Haze perception; Haze thresholds; Turbidimeter; Brightness model; Lightness model; Simultaneous lightness contrast; Digital imaging

# Acknowledgements

First, I would like to thank my supervisors: Associate Professor Dr. Roland Harrison, Dr. Karen Lusk and Professor Don Kulasiri. They have shown admirable patience and continuous support throughout.

I would also like to thank Richard Hider for his guidance in the lab, Dr. Sue Mason for her efforts in smoothing access to the sensory rooms, Bernard Newman for his help in the winery, and Tracey Gibson and Associate Professor Dr. Jim Morton for everything else.

The wine samples used in my experimental work were supplied by Villa Maria in Blenheim; thanks to Josh Hammond and Murray Cook for these.

Finally, thanks to all those unsung heroes who allowed me to investigate their haze detection talents (a special call out to those who came along to every experiment: Patricio Mejias Barrera, Dr. Glen Creasy, Dr. Carol Smith and Janette Busch).

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Research background

It is generally known that white wines may develop a protein haze during transport and storage; this is referred to as protein instability. Since haze in white wines is considered unacceptable by consumers, winemakers adopt various methods to mitigate the possibility of their wines becoming hazy prior to consumption. A heat test is commonly used to determine if the wine is susceptible to haze formation by examining protein stability. If the test classifies the wine as protein unstable, the winemaker is advised to adjust the levels of protein through the use of fining agents such as bentonite. An important step in heat test protocols involves the optimisation of bentonite dosage rates through the comparison of respective haze outcomes. While winemakers are advised to use a nephelometer to make the assessments[1], uptake of instrumentation is somewhat limited[2]. Instead, industry turns to the less costly visual assessment which is subjective and can be can be difficult to perform.

Research into the efficacy of fining practices suggests that the most commonly used heat tests can lead to excessive fining if the protocol recommendations for dosage rates are followed[2]. Winemakers are aware of the deficiencies in these tests and anecdotal evidence suggests that dosage reductions are commonly applied to test-derived fining rates as a means of mitigating product losses. The limited uptake of instrumentation as a means of assessing haze would suggest that outcomes are somewhat arbitrary in nature.

A further problem faced by the winemaker in gauging the relevance of haze levels is the absence of benchmark data relating instrumental measurements to human perception under realistic viewing conditions. Although some researchers have explored the relationship between human perception of haze and instrumental turbidity measurements[3–6], no such studies involving wine have been undertaken and haze perception thresholds in wines are yet to be reported.

The scale of economic loss due to fining depends on various factors but the best available estimate is around 3.5% by dollar value if rotary drum vacuum filtration is utilised for wine recovery[7]. Based on 2012 export statistics[8], the current white wine export value for New Zealand is around NZD 988 million; this would suggest a theoretical loss to the New Zealand wine industry in the order of NZD 35 million if the recommended protocols were being followed. Since winemakers are thought to self-regulate their fining levels, the true cost is likely to be somewhat less. However, every arbitrary

reduction in recommended dosage levels increases the probability of haze developing in the bottle. The true economic impact is therefore a combination of actual product loss due to fining combined with some level of commercial risk associated with the possibility of consumer dissatisfaction should haze develop.

## 1.2 Research goal and objectives

The overall goal of this project was to explore the relationship between nephelometry, digital imaging and human perception of haze in terms of thresholds and intensity. In addition, the feasibility of using the Android operating system as a platform for image based, nephelometric measurements was to be investigated.

Specific experimental objectives were:

- To assess the human visual response to haze intensity in polymer microsphere solutions and wine under various viewing conditions.
- To establish thresholds for haze detection in polymer microsphere solutions and wine under various viewing conditions.
- To investigate the relationship between nephelometric measurements and digital image data where the data is derived from images of polymer microsphere solutions and wine.

These experimental objectives were underpinned by a general wish to maintain relevance to real-world conditions; experiments were to be conducted using industry standard ISO glasses and commonly used illuminants.

## 1.3 Human ethics approval

All experimentation involving human subjects was conducted in accordance with Lincoln University Human Ethics Committee case approval: HEC 2013-47.

## 1.4 Thesis structure

This thesis comprises eight chapters including this general introduction. Chapter 2 reviews the literature related to wine protein stabilisation, turbidity measurements, scaling of haze perception, brightness and lightness psycho-physical models and haze detection thresholds. Chapter 3 provides details of common methodologies used throughout the experimental work. Chapters 4, 5, 6 and 7 are experimental sections encompassing haze detection, haze intensity perception and image based nephelometry in polymer microsphere suspensions and wine samples. Chapter 8 is a general summary including some comments on future work.

# Chapter 2

# Literature Review

## 2.1 Wine proteins and their stabilisation

### 2.1.1 Wine proteins

Protein stabilisation is a critical issue for white wines with many studies exploring their protein composition[9–13]. Most of this work has confirmed the presence of a small number of protein types including thaumatin-like proteins, osmotins, chitinases, invertases and β-glucanses ranging in molecular weight from 15 kDa to 66 kDa[14]. Some of the factors that appear to influence protein composition include cultivar, disease, harvest methods, harvest timing, water stress and grape variety[10,14].

The development of haze is normally attributed to unsuitable storage or transportation conditions (especially elevated temperatures) or extended storage times. Other factors that have been shown to affect haze development include the presence of compounds such as polyphenols, metal ions, sulphates and polysaccharides[15,16]. Adding to the overall complexity is the varying temperature sensitivity of the various protein fractions to heat-induced unfolding and precipitation[17]. In general, the mechanisms and kinetics of protein aggregation remain poorly understood[18] thus precluding the development of a predictive shelf life model.

### 2.1.2 Protein stability tests

Since the protein shelf life of a wine cannot be reliably forecast, winemakers who wish to ensure protein stability will work to reduce protein concentrations prior to bottling. In the absence of effective, predictive models the amount of fining agent required to make a wine stable is normally established through stability tests and fining trials. Over the years many empirically derived tests have been proposed[2]; all attempt to detect instability by artificially inducing precipitation either chemically and/or through heating. In a typical heat test a previously clarified wine sample is heated to an elevated temperature for a relatively short period of time (e.g. 80°C for 30 minutes[19]); the sample is then examined for turbidity after cooling. The outcome is a binary result: stable or unstable based upon a turbidity comparison of the heated sample versus the original. If the wine is deemed unstable, a range of fining agent additions are trialled and heat tested to determine the most appropriate dosage levels.

### 2.1.3 Wine industry methods of haze quantification

Whether the winemaker is conducting fining trials or long term storage tests, at some point they are faced with the need to assess the extent of haze in a sample. Although haze can be detected visually by passing a strong light across a wine sample, low level haze in the reference sample or colour development due to the stability test can make assessment difficult. Inconsistent results in visual assessments led Pocock and Rankine[2] to recommend the use of a nephelometer for haze comparisons. This advice aligns with current industry protocols recommending the use of a turbidity meter to objectively measure haze[1].

A further problem that arises from visual assessment of haze relates to the need to determine differential measurements. Most industry test methods cite a differential turbidity value that acts as a stability threshold when comparing samples (e.g. 2 Nephelometric Turbidity Units[1,13]). If the winemaker cannot arrive at an objective measure of turbidity, the method cannot be followed. Even if the winemaker is relying on past experience to judge acceptable haze levels, the lack of instrumental data prevents reliable use of historic norms.

## 2.2 Turbidity measurement in wine

### 2.2.1 Light scatter in hazy wine

As a beam of light passes through a liquid containing suspended particles, the light will be subject to both absorption and scatter. The scatter is caused by oscillatory interactions between the incident light and the particles resulting in secondary radiation leading to complex interference effects[20]. Critically, the intensity pattern of scattered radiation has an angular variation; this considerably complicates any attempt to characterise the scattering properties of a particular sample.

The general theory of light scattering for particles is known as the Mie model[21]; key variables include particle size, light wavelength, particle refractive index and scattering angle. This is a complex mathematical model with a number of limitations including an assumption that all particles are spherical. For particles much smaller than the wavelength of the incident light, Mie theory collapses to the simpler Rayleigh theory where the angular distribution of scatter is almost uniform. The total intensity of scattered light is generally categorised by the size of the particles in relation to the wavelength of the incident light[22]; summarised in Table 2.1. For particles in the order of the incident light wavelength (and smaller), a relatively small change in the particle size will have a correspondingly large impact on the amount of scattered light.

**Table 2.1 Total scattered light in proportion to particle size and wavelength of incident light for some constant number of particles.**

| Model | Boundary conditions | Total scattered light ($I_s$) |
|---|---|---|
| Rayleigh | $R \ll \lambda$ | $I_s \propto {R^6}/{\lambda^4}$ |
| Mie | $R \sim \lambda$ | $I_s \propto {R^4}/{\lambda^2}$ |
| Geometric optics | $R \gg \lambda$ | $I_s \propto R^2$ |

R is particle radius, λ is wavelength of incident light.

The physical nature (size, shape and refractive index) of haze inducing particles in wine is not well understood. In a study of heat effects in white wines Dufrechou et al.[15] found that exposing a sample of Sauvignon Blanc to 40˚C led to particle aggregation with an average hydrodynamic diameter of approximately 1.35 µm after a period of 2 hours. Any specific sample of wine undergoing protein unfolding and aggregation will contain an unknown distribution of particles with an associated (unknown) angular scatter characteristic. Figure 2.1 illustrates the relative scatter responses of a small selection of particle sizes (0.1 µm to 10.0 µm) based on Mie theory; note the large increases in light scatter for small increases in particle size along with the dynamic nature of the angular response. In a real-world sample of wine held within a standard drinking glass, the amount of scattered light (i.e. actual visual impact) will additionally depend on the absolute intensity of the incident light, optical characteristics of the glass and the volumetric dimensions of the wine[23].

### 2.2.2 Turbidity instrumentation

A turbidimeter operates by illuminating a test sample with an appropriate source of light and measuring the resulting scattered light. A conventional device will consist of a light source, sample holder and an optical detector set at some pre-determined location and angle. The simplest form of device is the single beam nephelometer which directs a light at the sample and seeks to detect the scattered light at a nominal angle of 90° generating measurements in Nephelometric Turbidity Units (NTU)[24].

There are a wide range of designs and test methods that vary by angle of detection, detector number and location, frequency and bandwidth of radiation source along with various other factors. The intention of these alternate design configurations is to minimize bias factors in a specific application setting[25]. Each of these instruments is measuring light at some specific angle(s) using light sources and detectors of a specific spectrum output and spectral response. This means the physical design of the instrument has a direct bearing on the final measurements; the devices are not independent of the attribute to be measured.

In an effort to derive standardised turbidity measurements, formazin (hydrazine sulphate and hexamethylenetetramine) is commonly used as calibration standard. Unfortunately design

differences in devices along with their discrete measurement angles means that measurements from different devices cannot be considered comparable despite the use of common calibration methods[26–28]. If true turbidity is taken to mean the total cumulative scatter associated with a particular sample taken over the full angular range, the complexity of the angular response in a material such as wine means that measurement at a specific angle cannot be used to infer overall turbidity[22]; formazin calibration does not resolve this issue[29]. Turbidity measurements of arbitrary materials are not traceable to a *relevant* physical standard; they are some device dependent measurement of appearance. Such measurements can only be considered relative and comparable to other measurements made on the same device type[30].

### 2.2.3 Relevance of turbidity measurements to wine stability thresholds

The traditional, non-instrumental approach to haze assessment is one based upon visual discrimination performed by the winemaker; this is a human perception task performed under ill-defined conditions. The use of turbidity instrumentation to replace the visual test is motivated by a desire to replace a subjective and variable approach with an objective and reliable measurement method. Methods that cite specific turbidity threshold values to declare a wine "protein stable"[1,31] assume that turbidity measurements are somehow absolute, comparable and relevant to perceptual levels of haze. Although the amount of haze perceived in a sample will generally increase with turbidity measurements, equivalent NTU values derived from different devices do not guarantee equivalent levels of visual haze[32].

## 2.3 Scaling of haze perception

Wine is a material that both transmits and scatters light such that its appearance is perceived to fall somewhere between the extremes of opacity and clarity[33,34]. As light transits a wine sample it is absorbed and scattered to varying degrees depending on the spectral nature of the light and the composition of the wine. As the concentration of haze causing material in wine changes, the proportion of scattered light to absorbed light adjusts accordingly. The perception of a hazy wine sample is a complex interplay of the physical attributes of the wine matrix, ambient lighting and the human visual system. Unfortunately little is known of the mechanisms employed by humans in the perception of such materials[35,36]; this presents a significant challenge in any attempt to scale the perception of haze.

### 2.3.1 Psycho-physical models of transparency and translucency

Hunter[37] has proposed a material appearance classification system based upon light transmission attributes whereby transparent materials are those that transmit light in a specular manner and translucent materials are those that transmit light in a diffuse manner. This is a simplification

designed for theoretical convenience and does not easily fit with the appearance of real-world materials such as hazy wine samples that may be both transparent and translucent.

The perception of transparency relates to the visibility of overlapping surfaces whereby the observer senses the stacked surfaces simultaneously. Since light reaching the eye does not contain discrete information pertaining to the different layers, the perception of transparency is the result of poorly understood neural processes[38]. Although there is a body of literature addressing the psycho-physical modelling of transparency[39–50], the dominant theoretical models are relatively simplistic in that they ignore many of the physical effects experienced under real viewing conditions. Important aspects not addressed include variable illumination, reflections between the observed material and other scene elements and gloss effects on the surface of the object.

In their investigations of translucency Fleming and Bülthoff[35] showed that many of the factors thought important to transparency were not required for translucency perception and that cues such as blur, shadow sharpness and colour saturation were more important. In a similar study Motoyoshi[51] showed that specular highlights and non-specular shading patterns can act as cues for the characteristic "glow" of translucent materials. In a recent study by Nagai et al.[52] the local mean luminance of specific image regions correlated well with perceptual translucency but the image regions contributing to the perception differed greatly between observers. Since all of these studies employed computer generated graphics as their basis for visual stimuli, their relevance to real-world objects and scenes is unclear.

There are currently no psycho-physical models that adequately describe the perception of real-world materials that simultaneously transmit and scatter light.

### 2.3.2 Empirical models of haze perception

Only a small number of empirical studies have sought to relate human perception of hazy liquids to instrumental measures. In one such study Malcolmson et al.[53] used Magnitude Estimation techniques to evaluate perceptual clarity for apple juice dosed with kaolin. The results were used to derive a log linear model between nephelometric measurements and perceived clarity. The resulting power function had a power exponent of -0.54 and a reasonably high $R^2$ value of 0.94 (Figure 2.2). The similarity of the exponent value to that found in studies of perceptual brightness was noted. Malcolmson et al. further claimed that a log transformed model could accurately predict the perception of clarity in apple juice over the examined range but a reconstruction of their data on non-log scales shows that the model fits poorly for the low turbidity range (Figure 2.3). This systematic deviance may have been due to induction effects (see Section 2.4) but the viewing conditions were not recorded preventing any further analysis.

Carrasco and Siebert[3] also used Magnitude Estimation to investigate the relationships between particle concentration, particle size, instrumental turbidity and visual haze intensity. In that study aqueous suspensions containing synthetic microspheres of various sizes and different pigment types (turbidity ranging up to 5000 NTU) were presented to human subjects under brightly lit conditions and dark backgrounds. The relationship between perceived haze intensity and microsphere concentration was found to be compressive whereas the relationship between instrumental turbidity and microsphere concentration was expansive. A 4$^{th}$ order polynomial relationship was derived using stepwise regression relating instrumental turbidity and perceived haze intensity across the entire dataset including all sizes and concentrations ($R^2$=0.87). A series of additional models were derived with 4$^{th}$ order polynomial combinations of particle size and instrumental turbidity offering the best fit for the data ($R^2$=0.978). Examination of a more detailed description of the base data[54] suggests that a power function model may have been fitted to the data if the concentration range had been truncated; a linear fit may be approximated on the log-log response up to around 300 NTU (see Figure 2.4). The experimental conditions employed in this study were tightly controlled involving the examination of samples inside a small viewing box with a dark background. The volume of the sample and the receptacle shape were not reported.

### 2.3.3 Haze as a sensory variable

An important factor in sensory studies is the definition of the perceptual variable; in this project's case *haze*. Carrasco and Siebert[3] performed sensory descriptive analysis to describe the visual perception "haze" in a model system. Applying Principal Component Analysis (PCA) to the data the authors found that two factors accounted for 99% of the variance. Factor 1 was aligned to a superimposed redundant cluster: "turbid" / "glowing" / "opaque" in one direction and "dark" in the other. Factor 2 was composed of two attributes: "particulate" and "homogenous". The particulate/homogenous axis only appeared active for the larger particle sizes (10.3 μm) in their study. These results suggest that it may be possible to use a scale of perceptual brightness as a surrogate for perceptual haze (at least where the particle sizes are small enough that the particulate dimension is not significant). Section 2.4 reviews brightness and lightness in terms of perceptual phenomena and Section 2.5 deals with the relevant scaling models.

## 2.4 Brightness and lightness phenomena

### 2.4.1 Overview

Any review of perceptual "brightness" and "lightness" cannot easily proceed without some elaboration of the terms. The most commonly used definition of brightness is stated as the attribute of a visual sensation according to which a given visual stimulus appears to be more or less intense, or according to which the area in which a visual stimulus is presented appears to emit more or less

light[55]. The corresponding definition of lightness is the attribute of a visual sensation according to which the area occupied by the visual stimulus appears to emit more or less light in proportion to that emitted by a similarly illuminated area perceived as a "white" stimulus[55]. It should be noted that these terms are entirely perceptual in nature and distinct from their physical counterparts of photometric luminance and surface reflectivity; measurements of photopic weighted light energy arriving at the eye or material reflectance do not correspond to perceived brightness or lightness. Although lightness and brightness are thought to be inter-related, little work has been published where they have been jointly considered in a real-world scene[56].

Although the definitions above appear to offer a sound basis for the study of lightness and brightness, individual experimental conditions and the exact nature of instructions given to participants has led to much confusion[57]. A recent review has shown that lightness judgments, brightness judgments and contrast evaluation are easily compounded or exchanged depending on the specifics of the study[58]. The authors cite a number of important studies that appear to be undermined by this issue. Finally, the psychological (and hence perceptual) relevance of the terms lightness and brightness has been questioned[47] despite their use in a vast number of studies.

Another set of terms that has relevance for this project is *unrelated/related colours*. In essence, unrelated colours are those of stimuli observed in isolation such as traffic lights on a dark night. Related colours are stimuli that are observed in the context of other stimuli[59,60]. Although it is not always a simple distinction to make, a visual stimulus in a sensory booth with a dark surround may be considered unrelated colour conditions. However, the inclusion of a second sample (e.g. a modulus reference) would likely convert the scene to related conditions. A sensory booth with a light surround would induce lightness effects and would be consistent with related colour conditions.

### 2.4.2 Lightness induction

The influence of surround conditions (taken to encompass both background and surround) on the perception of a target stimulus has been widely reported[61–65] with an extensive focus on simultaneous lightness contrast (i.e. lightness induction). The specific luminance of the surround is thought to be an important threshold that forms a boundary between two distinct perceptual responses known as *increments* and *decrements* whereby increment refers to conditions in which the luminance of the target exceeds that of the surround[66,67]. A second surround effect is related to enhanced brightness/lightness discrimination as the target luminance transits the luminance of the surround; this is known as the Crispening Effect[68–71]. Other attributes of scenes have also been shown to influence lightness appearance including relative areas[72,73] and planar grouping[74].

### 2.4.3 Chromatic effects

The level of colourfulness (saturation) in any stimulus will influence perceived brightness and this is known as the Helmholtz–Kohlrausch (H-K) Effect. Generally, this means that more saturated colours will appear brighter than those of low saturation even when they have the same measured luminance[75]. Yellow hues are generally thought to be less influenced by the H-K Effect but a study of unrelated colours failed to find statistical evidence to support this view[59].

Although much of this project was focused on the relationship between achromatic brightness/lightness and aggregate luminance (see Section 3.3.2), it is acknowledged that chromatic effects such as such as the H-K Effect and chromatic induction[76] will have some impact on the perceptual response.

## 2.5 Brightness and lightness scaling models

### 2.5.1 Overview

Many theoretical models have been proposed to explain the achromatic response of the human visual system. These theories tend to align with distinct philosophies and include edge integration, Gestalt anchoring, spatial filtering and layer decomposition models. A general theory that unifies the current theoretical fragmentation and predict the human response has yet to be formulated[77].

In the context of this project, the area of interest is best described by the everyday terms "brightness" or "glow" as perceived in real-world (albeit relatively simple) scenes. Whilst the debate over the distinction between brightness and lightness continues[47], it is not a simple matter to disentangle lightness from brightness in much of the literature. For example, the term brightness may be used to describe the perceptual response to a light source and at other times it may refer to an ill-defined, subjective correlate of luminance[78]. This confusion presents something of challenge in the selection of models that may have relevance and possible utility in the perception of haze. Based on an assumption of lightness constancy, Brill and Carter[79] suggest there are only two plausible mathematical models for lightness perception; in general forms:

$$y = a + b\log x \qquad \text{or} \qquad y = a + bx^k$$

For this reason, the models reviewed below are limited to those that employ relatively simple logarithmic or power functions. When reviewing the models, the terms brightness and lightness are used in a manner consistent with the studies being described.

### 2.5.2 Historical context

Attempts to construct a mathematical relationship between physical measures and the human perception of brightness have a long history. Kuehni[80] cites work by Aguilonius in 1613, Bouger[81] and Fechner[82]. Working from the basis of Weber's Law[83], Fechner proposed the perceived magnitude of a stimulus to be proportional to the logarithm of the physical intensity; a compressive, non-linear relationship[84]. Fechner's logarithmic relationship assumes that perceptual "just-noticeable difference" increments increase in proportion to the absolute intensity of the stimulus. Plateau[85] countered that ratios rather than differences were the basis for perception and that a power function best described the human response[65]. The log versus power function debate has continued unabated to the present day with little prospect of early resolution[79,80,86].

### 2.5.3 Power function models

In a series of studies by S.S. Stevens and co-workers, the basis for human sensory perception was systematically explored using the technique of Magnitude Estimation including the response to short duration achromatic flashes[87–90]. Stevens strongly advocated a power response for all the senses and proposed the relationship below[65]; this has become known as the Stevens Power Law[84].

$$\varphi = k(\emptyset - \emptyset_0)^n \qquad 2.1$$

In Equation 2.1 the dependent variable $\varphi$ represents the perceptual response, $\emptyset$ the physical stimulus, $\emptyset_0$ a threshold adjustment parameter associated with the physical stimulus, $k$ relates to the choice of units and $n$ is the exponent of the power response.

Stevens expounded the universality of this relationship in the human response and placed much emphasis on the exponent ($n$) publishing more than twenty values for various stimulus conditions[65]. Perceptual brightness was declared to have a compressive exponent of 0.33 in reference to a set of "standard conditions" (1 second duration flashes, dark adaptation, dark surround, 5˚ field of view). Although the influence of surround induction on the exponent was acknowledged (effects up to an order of magnitude), $k$ and $\emptyset_0$ were often referred to as constants alongside the prominent 0.33 exponent. In a later paper[91] the effect of adaptation was studied (still in dark surrounds) and increased levels of adaptive luminance were found to affect all three parameters. It was also found that once a subject's vision had fully adapted to the target luminance, the response appeared logarithmic over most of the range. These significant observations appear to limit the potential of Equation 2.1 to predict general brightness response in any real-world context.

### 2.5.4 Opponent induction models

In a study of chromatic contrast by Jameson and Hurvich[66], the experimental data showed that the brightness response (plotted in log-log coordinates) consisted of two distinct linear sections with an intersection where the surround and target luminance values were approximately equal. Jameson and Hurvich[92] subsequently proposed a joint chromatic and brightness model built on the premise that simultaneous contrast effects are due to brighter surround areas inducing an opposite, physiological, blackness response primarily associated with the target area. Horeman[93] reviewed a number of studies and also showed that the brightness response in the presence of surround induction appeared to form dual segments when plotted on a log-log scales; Figure 2.5 shows a typical brightness response for varying levels of induction.

Combining lateral inhibition with a number of other physiological effects, Haubner et al.[94] developed a model that extended the work by Jameson and Hurvich to include field of view parameters; its general form is shown below:

$$B = C_T(\varphi)L_T^n - B_0(L_u, \varphi) \qquad 2.2$$

$$\text{where} \qquad B_0 = C_T(\varphi)[S_0(\varphi) + S_1(\varphi)\, L_u^n\,] \qquad 2.3$$

For a target area $L_T$ subtended by a field of view of $\varphi$ with a surround luminance of $L_u$ the brightness is given by $B$. The exponent $n$ is fixed at 0.31. $C_T$, $S_0$, and $S_1$ are all constants dependent on the angle of view; these are given in a table of constants by Haubner et al. Bodmann[95] claimed that the various logarithmic and power laws derived by other workers are simply sectional approximations of this model. A key difference from the Stevens Power Law is the explicit handling of the surround luminance whilst maintaining a constant exponent value for both target and surround terms. The model does not appear to accommodate the Crispening Effect and the effect is not readily visible in Haubner et al.'s experimental data (see Figure 2.6).

### 2.5.5 Log W

In the tradition of Fechner and Weber, Whittle[70] had subjects set an equal interval scale using 25 circular shapes on a computer display to represent the full greyscale range from black to white in a mid-grey surround. Whittle's data (Figure 2.7) captures two key aspects of simultaneous contrast in simple patch scenes: (1) a distinct composite nature (2) a steepening response at the intersection (the Crispening Effect). This is shown in Figure 2.7 which also illustrates the poor fit of Stevens Power Law to the data. Whittle proposed a logarithmic function (Log W) to predict brightness contrast

response under both incremental and decremental conditions. Working in the context of display technology Carter[96] elaborated on Whittle's original work to propose two specific grey-scale functions; one for increments and one for decrements. Although Carter provides detailed justification for the various parameters involved, many are approximations that depend on the specific viewing conditions.

$$B^{+} = 8.22\ log_{10}\left(1 + \frac{\Delta L\left(1 - k_p\right)}{0.15\left(L_b + L_{dk} + \ k_p \Delta L\right)}\right) \quad 2.4$$

$$B^{-} = -7.07\ log_{10}\left(1 + \frac{\Delta L(1 - k_n)}{0.15(L_a + L_{dk} + \ k_n \Delta L)}\right) \quad 2.5$$

$B^{+}$, $B^{-}$ represent brightness contrast in incremental and decremental conditions respectively, $\Delta L$ the absolute difference of target to surround luminances, $L_b$ luminance of the surround, $L_a$ luminance of the target, $k_p$ and $k_n$ ocular scatter adjustments and $L_{dk}$ an estimate of dark light levels in the eye. Note the model is constructed to predict relative brightness contrast; as $\Delta L$ tends to zero, the function predicts zero brightness contrast. Carter[96] lists a number of limitations and cautions the model may not apply to three dimensional objects. Despite these deficiencies the model is comparatively simple, predicts the Crispening Effect and accommodates scenes containing both dark and a bright surrounds.

### 2.5.6 Colorimetric brightness/lightness models

The science of colorimetry deals with the practical measurement of colour and as such has a long history of trying to accommodate brightness and lightness effects in its models. This has led to the widespread use of cubic power functions in colour coordinate systems and colour spaces such as CIE XYZ and CIELAB[84] that make no allowance for surround induction effects. Bartleson and Breneman[97] studied the effects of surrounds on printed images and proposed power exponents of 0.33 for dark surrounds, 0.41 for dim and 0.50 for light surrounds[98]. Modern colour appearance models such as CIECAM02 use a modified cube root function with parametric values for broadly defined conditions of average, dark and dim surrounds[99]. The underlying model of CIECAM02 has been shown to poorly predict the effect of surround contrast[100]. Wu and Wardman[101] have proposed modifications to CIECAM02 to improve its performance (CIECAM-m2) but they caution that their optimized parametric values only apply to the experimental viewing conditions they employed. All of the CIE lightness models are dependent on the Y coordinate of the XYZ tristimulus values and these require a properly characterised colorimetric instrument for measurement. The difficulty of capturing XYZ tristimulus values prevented the use of CIECAM02 in this project.

### 2.5.7 Real-world ambiguities

Even in relatively simple scenes the influence of induction can be substantial and surround conditions may radically shift the perception of target areas if the surround is of a higher luminance. Marsden[102] provides a useful illustration of the likely relationship between luminance of a target and perceived brightness in a complex scene (Figure 2.8). In an isolated dark surround the line QP would represent the response with a slope of around 0.3 (modified slightly by differences in chromatic content). The introduction of surround (or adjacent) elements will instigate both induction and lightness effects. Induction will cause the brightness of the target to drop and the simple line of PQ is transformed to a curvilinear triangular area of PRS. Lightness effects under general illumination will modify the response further by increasing high luminance elements and lowering the low luminance elements producing a final locus of TPVR. According to Marsden the actual response in a real scene is located in the TPVR envelope in some disordered manner.

## 2.6 Haze detection thresholds

A number of studies have sought to establish haze detection thresholds in polymer microsphere solutions using different surrounds and levels of illumination. Carrasco and Siebert[3] employed clear, yellow and red suspensions in a dark surround and brightly illuminated viewing box; thresholds were found to range from 0.38 to 0.81 NTU. Fleet and Siebert[5] extended this study by varying the levels of illumination and found that an intermediate lighting level (592 Lux) resulted in the lowest thresholds for clear, tan and brown suspensions; 0.176 to 0.829 NTU. Higher and lower levels of illumination (1192 Lux, 18 Lux) led to an increase the observed thresholds. Fleet and Siebert[6] used the same viewing box in conjunction with bright illumination (1192 Lux) to explore the effect of varying the surround conditions. Black velvet surrounds resulted thresholds ranging from 0.21 to 2.2 NTU whereas white cotton surrounds led to a much higher range: 2.0 to 41.0 NTU. The sample dimensions were not reported.

Horne et al.[4] attempted to derive haze detection thresholds for model solutions viewed under conditions that approached point of sale conditions. In this study the models suspensions were created by mixing water, formazin and red/yellow coloured pigments so as to emulate apple juice. The subjects were permitted to handle the test samples (presented in cuvets, 95mm x 25mm diameter) in a viewing room illuminated by both fluorescent and incandescent bulbs. Mean threshold levels were found to be much larger than those found using a tightly controlled viewing box: 3.6 NTU. Although the study was intended to emulate “store like” conditions, the levels of illumination and surround conditions were not specified thus limiting the usefulness and relevance of the threshold values found.

Horne et al.[4] compared the use of Ascending Method of Limits (AML)[103] with a transformed staircase method for the determination of haze thresholds; they found equivalent results for the two methods and recommended AML due to its comparative simplicity. All of the other studies employed the AML method to determine best estimate group thresholds.

## 2.7 Literature summary

Current industry protocols for the stabilisation of proteins in wine depend on the assessment of visual haze outcomes as a means of setting dosage levels for fining agents. The use of formazin calibrated nephelometry has been proposed as surrogate instrumental method for visual haze assessment. Nephelometers offer a device specific assessment of a light interaction with a liquid sample. The mechanisms used by the human visual system in the perception of translucent liquids such as wine are not well understood but it appears that many factors relating to the visual scene will influence the outcome. There is currently no scaling model available that can be used to estimate haze perception intensities based upon nephelometric measurements.

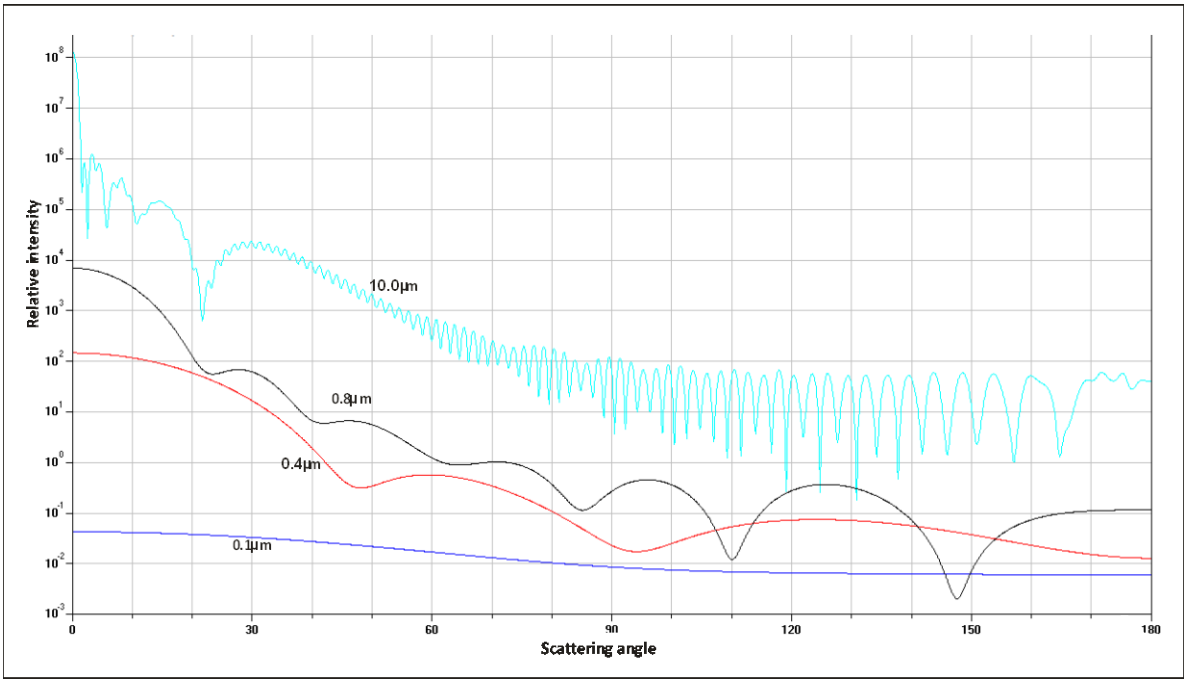


**Figure 2.1** **Relative scatter light intensities versus angle for particles of varying sizes (0.1µm to 10µm, unpolarised, λ=570nm). Derived using MiePlot version 4305[104].**

Figure removed due to copyright restrictions.

**Figure 2.2** **Relationship between apple juice clarity and instrumental turbidity. Adapted from Fig 1 in Malcolmson et al.[53].**

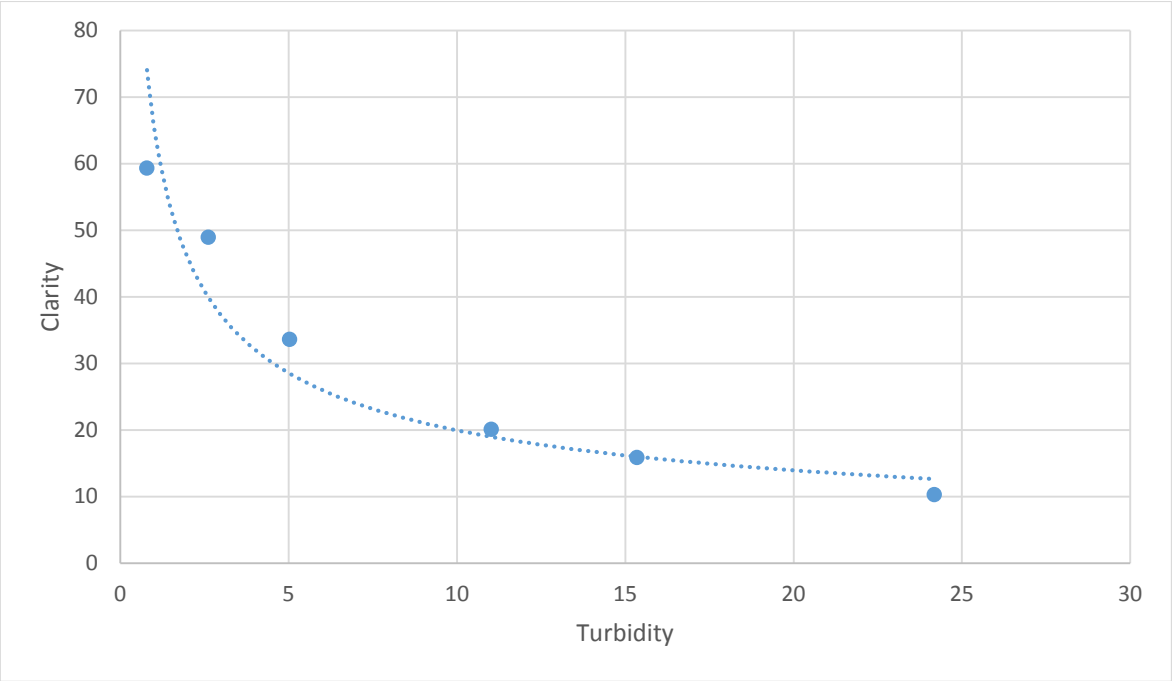


**Figure 2.3** **Relationship between apple juice clarity and instrumental turbidity. Approximated from Fig 1 in Malcolmson et al.[53]. Reconstructed data shown with a best fit power function of exponent value of -0.52.**

Figure removed due to copyright restrictions.

**Figure 2.4** **Haze intensity as perceived by sensory panellists versus turbidimeter for clear, yellow and red samples. Data for 0.769µm microspheres on left, 2.6µm on right. Below 300 NTU both sets of data are approximately linear. Illustrations sourced and adapted from Fig 3.7 and Fig 3.8 in Carrasco[54].**

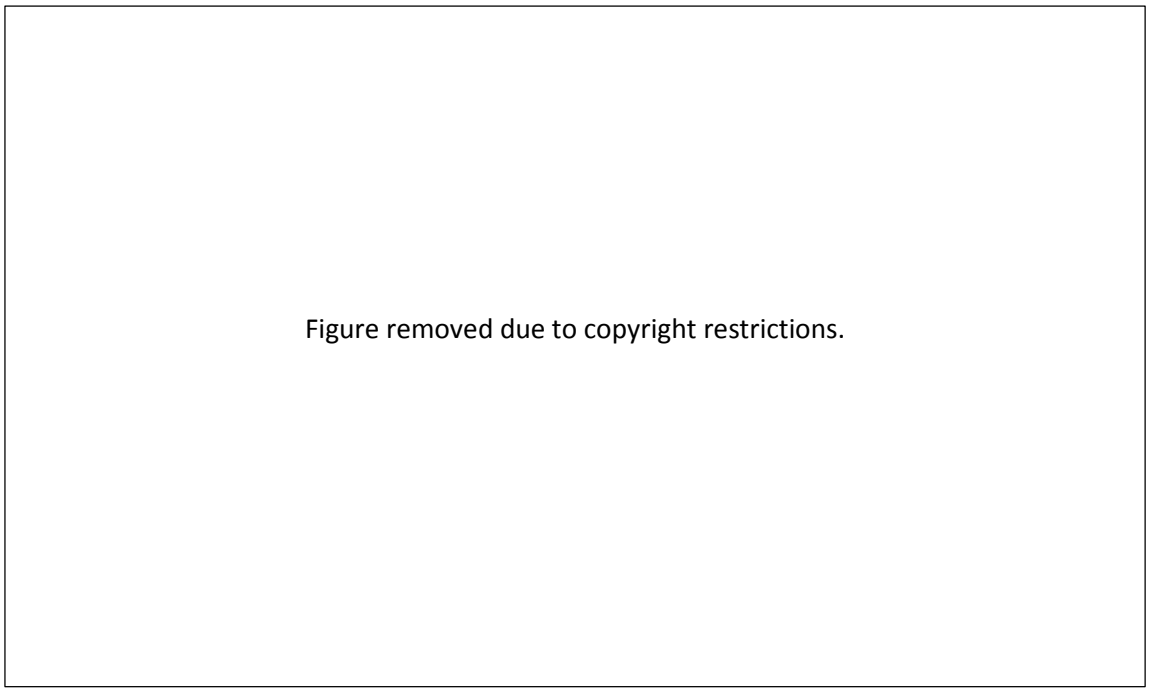


**Figure 2.5** **An example of dual segment brightness responses versus target luminance under varying levels of induction. The upper line represents the response in the absence of induction. The various steeper segments are observed at the levels of induction marked on the segments. Adapted from Fig. 4 in Horeman[93].**

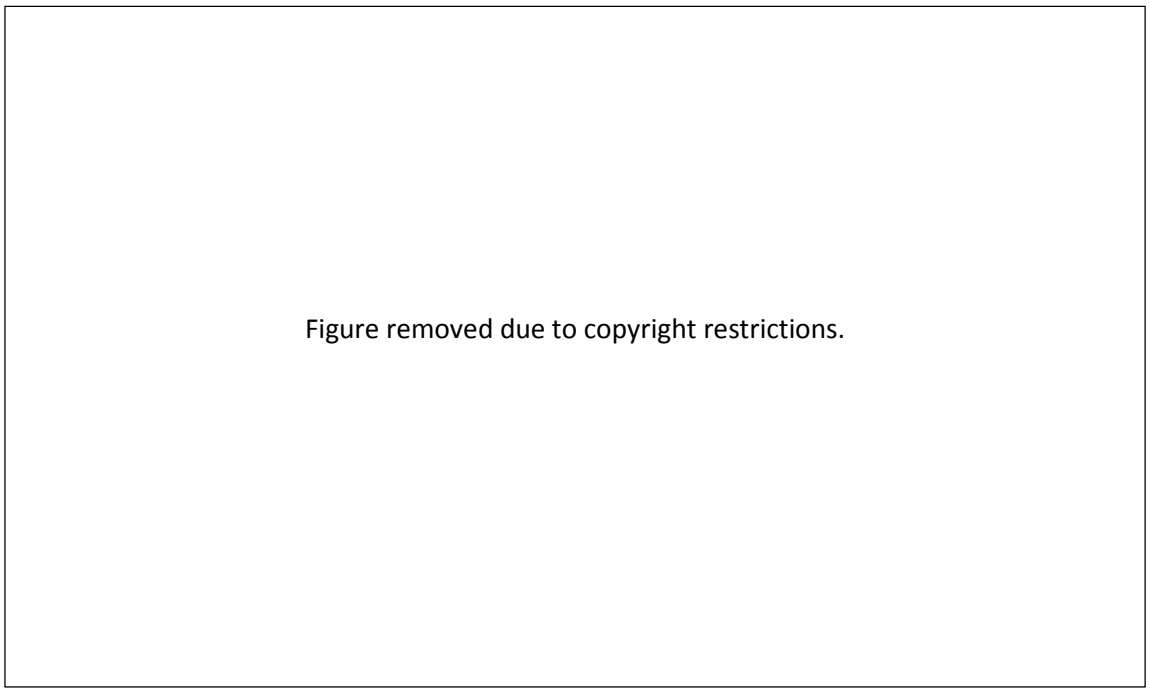


**Figure 2.6** **Achromatic brightness scaling data adapted from Fig. 6 in Bodmann[95]. The upper linear segment data corresponds to a dark surround, the lower curve data corresponds to a surround luminance of 300 cd.m$^{-2}$. Solid lines represent Haubner's model (see Equation 2.2 in this document).**

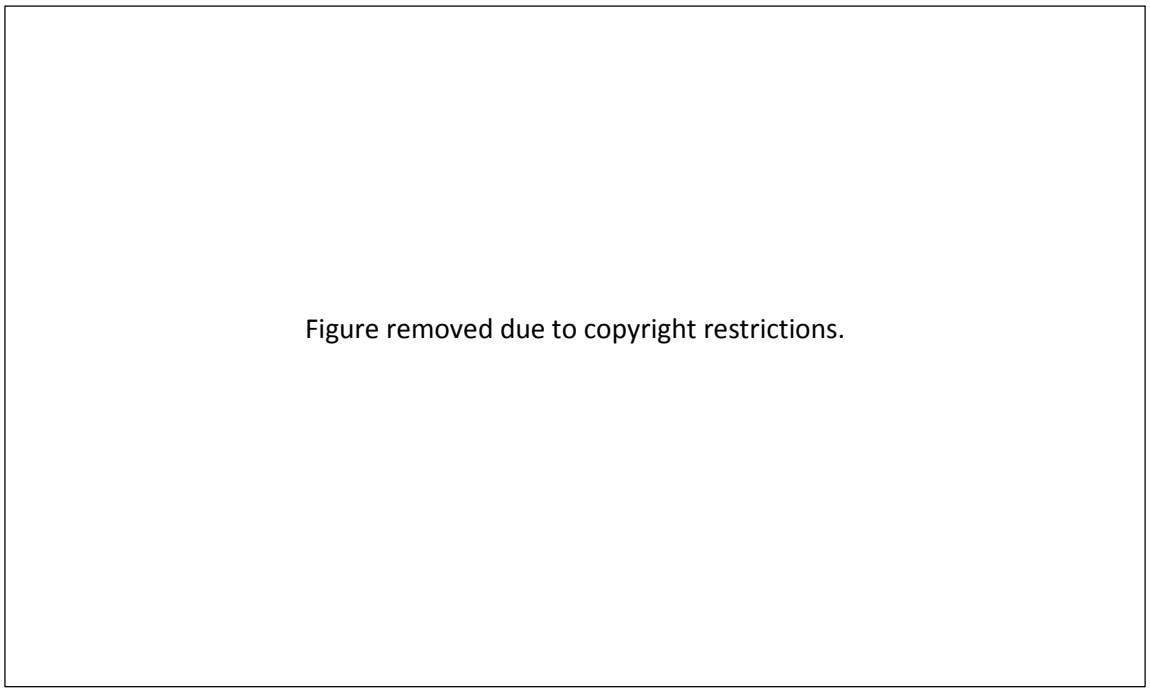


**Figure 2.7 An equal interval brightness scale for 25 grey circles on a grey background, plotted against luminance and log luminance. The ordinate axis is the ordinal number of the circles with the origin shifted to set brightness to zero at the position where the background luminance ($L_b$) intersects the curves. The dashed line is the best fitting Stevens power law with an exponent of 0.46. Adapted from Fig. 3 in Whittle[70].**

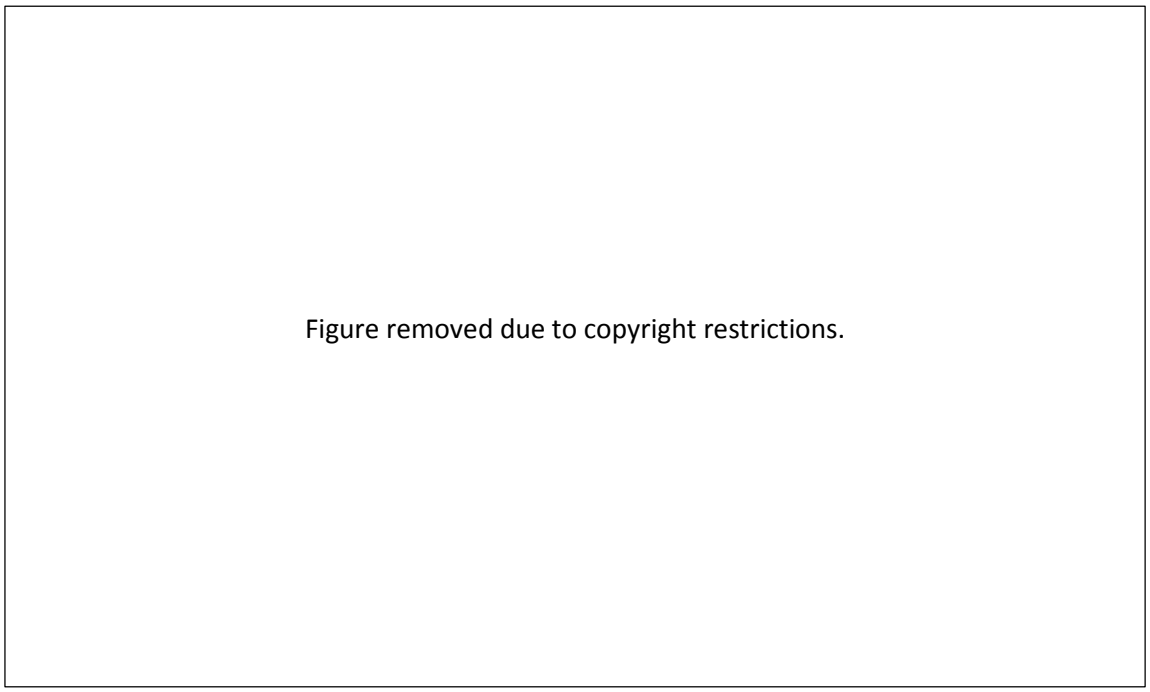


**Figure 2.8 Estimated locus of brightness (B) versus luminance (L) for varying conditions. TPVR represents an approximation to a real-world scene. Adapted from Fig. 11 in Marsden[102]. Refer to text for description (Section 2.5.7).**

# Chapter 3

# General Methodology

## 3.1 Overview

This thesis is composed of a number of inter-related studies that utilise many common methods and materials; this section details generic elements of the work. Additional information regarding methods and materials may be found in the relevant experimental sections.

## 3.2 Sensory environment

### 3.2.1 Sensory booths

In the experiments involving human participants, the perceptual response to haze was investigated in a sensory room composed of individual booth accommodation. The dimensions of the booths were 2.7 m high, 0.8 m wide and 1.1 m deep with a built-in table top at a height of 0.72 m. Each booth had its own illumination provided from a triple bulb enclosure at a height of 2.0 m centrally located over the table top.

The sensory room contained eight sensory booths of which four were modified to match two surround configurations. This resulted in two dark surround booths and two light surround booths (sufficient to support the anticipated maximum number of experimental configurations). This was achieved by hanging dark or light linen material from a height of 1.7 m; the linen covered the rear and side walls. In addition, a flat base covered with appropriately coloured material was placed on top of the built-in table top. The modifications were such that the subjects' field of view was completely enclosed by the dark or light material when seated in the booths. One of the sensory booths equipped with a dark surround is shown in Figure 3.1.

### 3.2.2 Booth illumination

The standard illumination was modified in the booths by replacing the existing bulbs with two LED bulbs (model: LED-MR16-SP-CW; Marexim Ltd., Mt Maunganu, New Zealand) in each of the four experimental booths. The bulbs were installed in the two outer positions in a triple bulb housing in an attempt to achieve a symmetrical illumination across the viewing area in the booth. The bulb housing did not have a diffuser installed meaning that the illumination pattern peaked in the central area with an increasing drop off towards the side of the booths. The bulbs had a claimed luminous flux output of 400 lumens, colour appearance of 6400 Kelvin and a colour rendering index of 80. The

spectral power distribution curve for the LED-MR16-SP-CW bulb exhibits a narrow maximum peak at 450 nm and a broader secondary peak at 550 nm.

The resulting illumination in the booths ranged from around 2000 Lux at 1.2 m above floor level (subject eye level) to around 700 Lux at 0.72 m above floor level (table top). These illumination values conform reasonably well to the range of 750 to 1200 Lux recommended for sensory room conditions by the American Society for Testing and Materials (ASTM)[105] ensuring that photopic conditions prevailed[106].

### 3.2.3 Sample receptacles and sample volume

All samples were evaluated by subjects using elongated egg shaped ISO glasses[107] typically used for wine evaluations. The glasses were manufactured by Arcoroc International (Arques, France) with all glasses originating from the same manufacturing batch. The physical dimensions of the glasses were: opening diameter of 4.2 cm; cup diameter (maximum diameter of bowl) of 6.0 cm and a cup height of 9.0 cm. The optical transmission characteristics of the glass material were unknown.

All human subject evaluations were conducted with sample volumes of 90 mL.

### 3.2.4 General viewing conditions

Although viewing conditions were generally dictated by the booth environments, the subjects were encouraged to examine the samples in a natural manner with a requirement to commence evaluation by first viewing the sample at eye level. Prior to each evaluation session the subjects' chairs were adjusted so that eye level approximately aligned to 510 mm above the booth table top. Assuming an eye-to-glass distance of 35 cm, this equates to a combined stimulus/proximal field of approximately 10°. The background and surround conditions were dictated by the homogenous linen material hung in the booths combined with the spatial characteristics of the illumination.

Subjects were specifically requested to hold the glasses by the stem so as to avoid contamination of the viewing surface.

The choice of achromatic material for the booths was intended to minimise any potential chromatic induction effects. This was largely achieved with the dark surrounds but the response of the light surround had a bias towards green. Figure 3.2 shows device RGB values for each surround type in conjunction with the booth illumination (see Section 3.3.2 for measurement methodology).

## 3.3 Instrumentation

### 3.3.1 Nephelometric measurement

Turbidity measurements were conducted using an industry standard nephelometer (model: 2100P, Hach Pacific, New Zealand) conforming to USEPA Method 108.1[108]. This instrument functions by calculating turbidity from the ratio of scattered light at 90˚ to transmitted light using a tungsten filament light source. Measurements are displayed in Nephelometric Turbidity Units (NTU) referenced against a formazin (hydrazine sulphate and hexamethylenetetramine) calibration.

Measurements were made by filling a sample cell provided with the instrument with a 15 mL volume of the sample with the *averaging* and *auto-ranging* functions of the instrument active. Where randomisation did not dictate otherwise, sequential measurements were made in increasing order of expected NTU values. Prior to each measurement the sample cell was rinsed with the liquid under test.

Calibration was performed against the *StablCal* stabilized formazin standards provided by the manufacturer prior to experimental work commencing. Calibration checks were made against secondary *Gelex* standards as recommended by the manufacturer.

### 3.3.2 Image based aggregate luminance measurements

Image based aggregate luminance measurements were conducted using a modified Canon A2300 digital still camera in conjunction with a number of supporting software components. An overview of the image processing pipeline is illustrated in Figure 3.3. Specific details of the camera setup along with software parameters may be found in Appendix A.

The sensor values retrieved from the camera can be significantly affected by the choice of image areas used for analysis. Pilot work was conducted to identify the most appropriate the region of the ISO glass to use as the Region of Interest (ROI) for sample analysis. It was found that many areas of the glass were affected by reflections from surrounding elements such as booth walls along with various specular glare effects. These interferences were minimised when the ROI was set to an upper central region in the glass for an orthogonal image capture geometry. This ROI and geometry was used in all studies (see Figure 3.4) except where the sample surface was imaged directly from above (see Chapter 7).

Note that *aggregate luminance* is used in this thesis as a term representing a weighted average measure of the combined Bayer colour filter array response from the camera as presented at the output of the processing pipeline; it is not intended to represent an absolute radiometric or CIE photometric luminance measure[109]. The weighting scheme used to average the RGB channel values:

$$\text{Aggregate luminance} = 0.2126\,R + 0.7152\,G + 0.0722\,B \qquad 3.1$$

R, G and B represent the raw channel, device dependent responses received at the output of the image processing pipeline shown in Figure 3.3. The weights chosen are arbitrary since the spectral response of the device was unknown but they do conform to ITU-R BT.709 luminance for linear primaries[110]. Aggregate luminance data is reported in relative form typically normalised against the highest luminance sample in the dataset.

### 3.3.3 Illuminance measurements

Illuminance measurements were conducted with a photography and lighting quality illuminance meter (Model: Digitech QM1587; Jaycar Electronics, Auckland, New Zealand). This instrument displays measurements in Lux corresponding to the CIE V($\lambda$) photopic luminous efficiency curve[111].

## 3.4 Reagents and sample preparation

### 3.4.1 Reagents and materials

Polystyrene, non-functionalised microspheres of mean diameter 0.25 µm were sourced from Bangs Laboratories Inc. (Fishers, IN, USA); product code: PS02N/7307. The microspheres had a dry density of 1.05 g/cm$^3$, an unspecified standard deviation and were supplied in an aqueous solution composed of de-ionised water, 0.1% sodium dodecyl sulphate (SDS) and 0.05% sodium azide ($NaN_3$). The number of microspheres per mL was calculated by the manufacturer to be 1.228x10$^{13}$. The microspheres were received in small plastic bottles with an approximate volume of 5 mL.

Yellow food colouring concentrate distributed by Hansells Food Group (Auckland, New Zealand) was sourced locally. The colouring concentrate contained water, tatrazine, carmoisine, citric acid, potassium sorbate and sodium benzoate (2.3% total dyestuff).

Sodium dodecyl sulphate, *specially pure1* grade (VWR International Ltd., Poole, UK) was purchased locally and used as a surfactant in the model microsphere solutions.

Visking dialysis tubing of molecular cut-off weight of 12-14000 Da (sizes 18/32 and 36/32) was supplied by Medicell International Ltd, London, UK.

Chardonnay wine samples were vinified by Villa Maria Estate (Marlborough, New Zealand) from grapes grown in the 2012 season. The wine was stored at 4˚C prior to and during the experimental period.

### 3.4.2 Preparation of model microsphere suspensions

Microsphere solutions were prepared at various concentrations through the dilution of microsphere concentrates using HPLC grade deionised water containing SDS. Preparation of samples involved a two stage dilution process whereby base concentrates were first prepared for use in a second stage dilution used to create a range of final sample concentrations. Dialysis to remove sodium azide was conducted before the second stage dilution (see Section 3.4.4). Prior to commencing each stage of the dilution process the relevant microsphere concentrate was rotated overnight in a Stuart incubator (model: SI30H, Bibby Scientific Ltd., Stone, UK) to ensure the microspheres were well dispersed. During pilot work it was established that maintaining an SDS surfactant concentration of 0.1% in the final samples would induce foaming effects that were likely to interfere with turbidity measurements and visual assessments. The final stage dilution therefore involved the reduction of the SDS concentration to 0.005% (a level previously used by the manufacturer).

The density of the microspheres means they approximate neutral buoyancy in water and should therefore create stable aqueous suspensions. Pilot work showed that suspensions remained stable for at least a week after preparation rendering density modification of diluents unnecessary.

Pilot experimentation using visual matching in one of the experimental dark surrounds indicated that 0.025 mL of yellow concentrate in 1000 mL of aqueous solution (SDS, 0.005%) would result in a solution with visual characteristics close to that of a locally available commercial Riesling wine (Hardys Riesling, 2103). All coloured microsphere suspensions were prepared to this pigment concentration by adding yellow concentrate to the diluent at the second stage dilution.

The volumes and dilutions used were specific to the relevant experiment; details can be found in the experimental sections below.

Microsphere concentrates and sample solutions were stored at 4˚C once prepared. Samples were removed from refrigeration and allowed to reach room temperature prior to experimental use.

### 3.4.3 Preparation of haze induced wine samples and stability assurance

Haze induced wine samples were prepared by heating commercial grade wine stocks at a high temperature and then using this material as a concentrate for dilution with unheated wine; the specific heating protocol and dilutions can be found in the relevant experimental sections.

Various studies had reported difficulties in working with natural haze material due to its tendency to be somewhat water soluble leading to general instability in experimental samples[5,112]. This is further compounded by heat and time related protein unfolding and aggregation processes that are inherent to the wine matrix[17]. Pilot work showed that haze induced wine samples prepared by “heat and

dilute" processes were reasonably stable over the course of a day if the turbidity was sufficiently low (Standard Deviation < 0.15 over the course of eight hours below 10 NTU).

To avoid excessive error due to sample evolution, all experimental studies involving haze induced wine samples were conducted over the course of a single day. Where samples in excess of 10 NTU were used, multiple turbidity measurements were taken at regular intervals in order to better estimate mean turbidity values and associated error.

### 3.4.4 Dialysis of microsphere concentrate

Prior to use with human participants, the sodium azide component of the microsphere concentrate was removed by dialysis. Suitable lengths of Visking's dialysis tubing were prepared by soaking in running water for 30 minutes followed by rinsing both the exterior and interior with lab grade reverse osmosis water. An aqueous solution of 0.1% SDS was prepared using HPLC grade water for use as the dialysis buffer. Dialysis was conducted as a three stage process using a fresh 2.5 L volume of buffer for each stage (the volume of concentrate to be dialysed varied by experiment).

## 3.5 Statistical and analytical methods

### 3.5.1 Inter-subject variance analysis

Inter-subject variance in sensory experimental data was assessed by calculating a coefficient of variation (CV) value for each subject and comparing it with the geometric mean for all subjects in the manner of Withouck[59]:

$$CV = 100\sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\frac{\left(Q_{geom,i} - fQ_{subj,i}\right)^2}{\bar{Q}_{geom}^2}\right)} \qquad 3.2$$

with $$f = \frac{\sum_{i=1}^{n} Q_{geom,i}Q_{subj,i}}{\sum_{i=1}^{n} Q_{subj,i}^2}$$

$Q_{subj,i}$ represents the haze perceived by the subject for the test sample i, $Q_{geom,i}$ the geometric mean of the $Q_{subj,i}$ for all the subjects, $\bar{Q}_{geom}$ the arithmetic mean of $Q_{geom,i}$ for all test samples, *n* the number of test samples, and $f$ the factor adjusting the $Q_{geom,i}$ and $Q_{subj,i}$to the same scale.

### 3.5.2 Curve fitting and regression analysis of haze estimation data

Curve fitting and non-linear regression analysis was conducted using CurveExpert Professional software[113] which uses the Levenberg-Marquardt method; six custom models were defined for this project (see below). The same software was used for linear fitting.

In a study by Marks and Stevens[114] a number of brightness scaling power law variants intended to accommodate surround induction effects or deal with observed deviances close to the perceptual threshold were reviewed. These models were elaborated into general forms for use as custom models for non-linear regression:

$$y = k(x - b)^n \tag{3.3}$$

$$y = k(x^n - b^n) \tag{3.4}$$

$$y = k[(x + b)^n - b^n] \tag{3.5}$$

In addition, the following logarithmic formulae were also used as base models for non-linear regression. Equations 3.7 and 3.8 are generalisations of the Log W model outlined in Section 2.5.5.

$$y = \mathrm{k}\ln(x) \tag{3.6}$$

$$y = k\log(1 + d(x - b)/b) \tag{3.7}$$

$$y = k\log(1 + d(b - x)/x) \tag{3.8}$$

The independent variable $x$ represents a physical measure (normally device dependent aggregate luminance or nephelometric turbidity). It may also represent alternative physical measures such as instrumental turbidity or microsphere concentrations. The dependent variable $y$ represents the haze sensory response corresponding to the experimental evaluations obtained through sensory studies.

The exponent $n$ is the power function value that influences the form of the power response. The parameter $b$ is an offset that may or may not relate to some physical measurement such as surround luminance or perceptual thresholds. The parameter $k$ is a scaling factor that allows the independent variable data to align to the arbitrary sensory scale. The multiplicative parameter $d$ is required in Equations 3.7 and 3.8 in order to conform to the general structure of the empirical Log W model.

The approach adopted for curve fitting was to sequentially attempt to fit the data for each general model and then examine the residual runs and associated p-values. The various parameters were examined for plausibility and a final visual check of the resulting curves were made for systematic deviances. The optimum logarithmic model and optimum power model were then selected for summary presentation in the relevant results.

**Figure 3.1 Sensory booth in a dark surround configuration.**

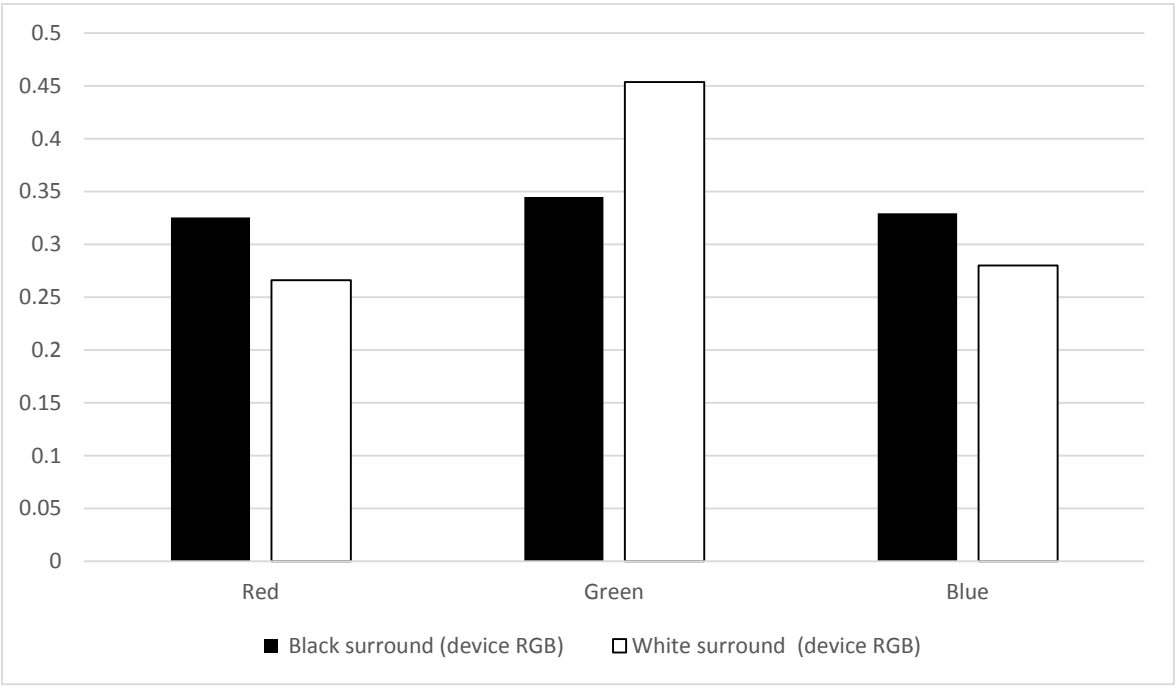


**Figure 3.2 Relative compositions or red, green and blue (RGB) channels in light and dark surrounds (based on device RGB proportions).**

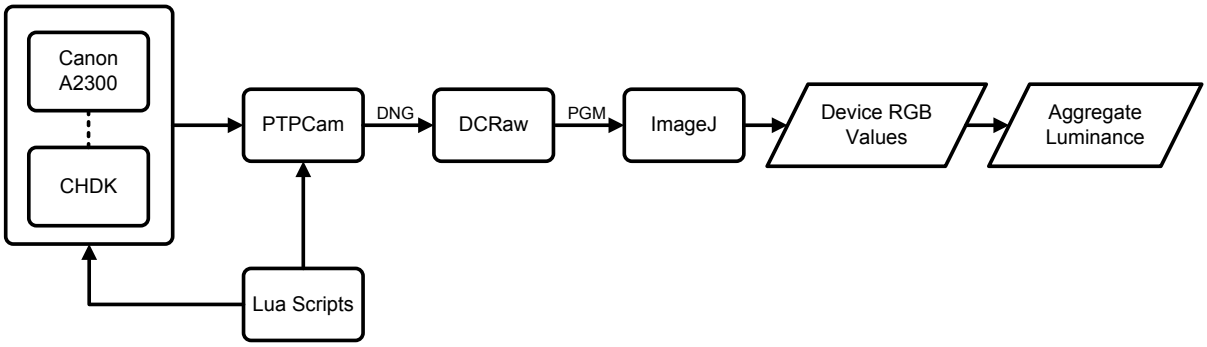


**Figure 3.3** **Image processing pipeline (see Appendix A for detailed explanation).**

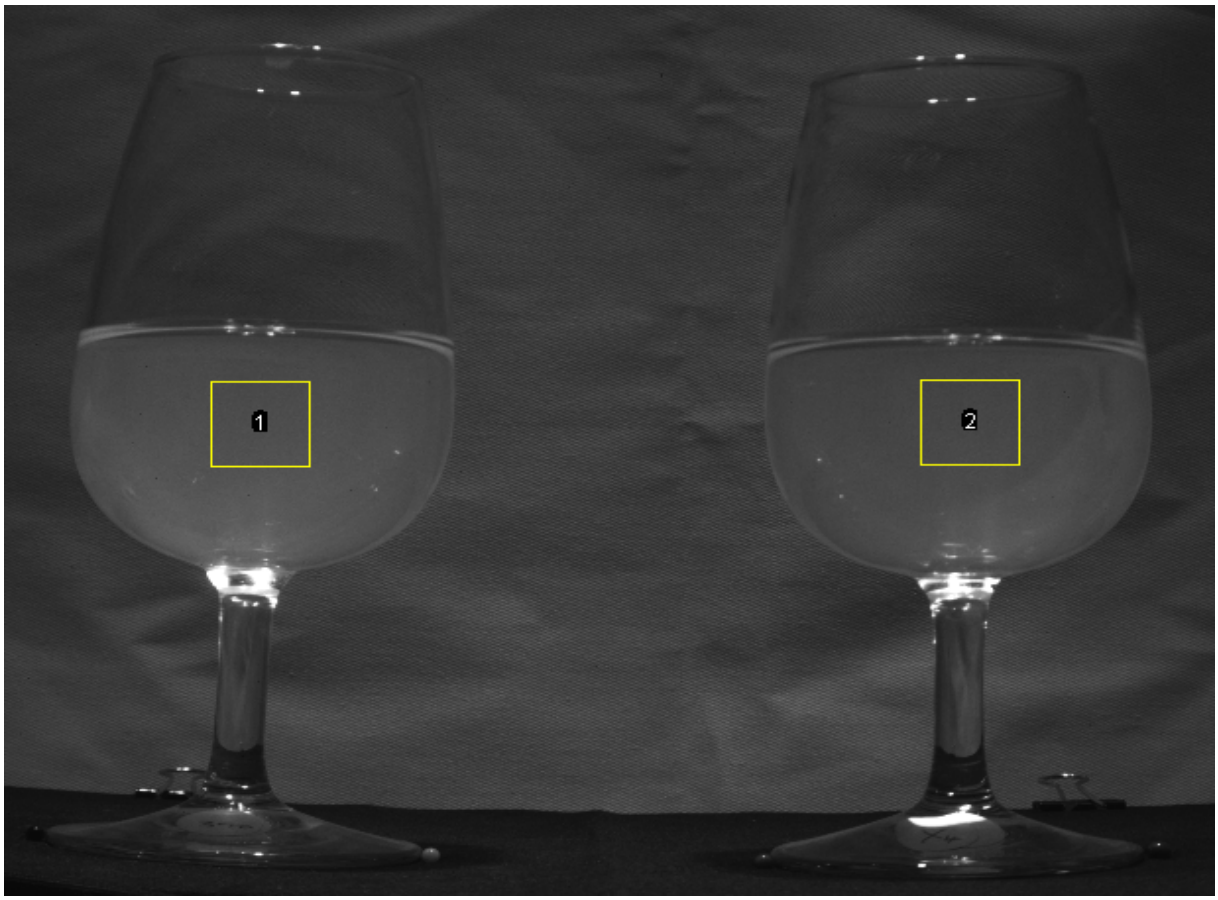


**Figure 3.4** **Upper central region of interest (ROI) areas used in image capture. Various areas of reflections and specular highlights visible outside the ROI's.**

# Chapter 4

# Haze Detection Thresholds in Polymer Microsphere Solutions and Haze Induced Chardonnay Wine Samples

## 4.1 Overview and objectives

Of the small number of studies that have examined haze detection thresholds, the majority have chosen to use microsphere suspensions of varying sizes and colours[3,5,6]. These studies have employed highly controlled conditions using a small viewing box to hold the samples static during evaluation. However, wine is consumed under very different conditions whereby the liquid is poured into a glass which the consumer is free to manipulate prior to and during consumption. The visual appearance of a wine in a glass is not homogenous and varies as the glass is moved relative to lighting and surrounding conditions. Therefore viewing arrangements that prevent subjects from handling test samples may result in detection thresholds that lack relevance to realistic conditions of consumption.

The overall goal of these experiments was to investigate haze detection thresholds in samples viewed under conditions resembling those encountered when wine is being consumed. Test materials were to consist of polymer microsphere solutions along with commercial grade Chardonnay wine samples.

The primary objective was to determine if different surround conditions (dark or light) or different coloured samples (achromatic or yellow) result in different haze detection thresholds. A secondary objective was to determine if haze detection thresholds in wine differed from those in equivalent yellow polymer microsphere samples under dark surround conditions.

## 4.2 Materials and methods

The experiment was executed in two stages: (1) threshold determinations using a range of surround conditions and suspension colours using polymer microspheres of 0.25 μm diameter (2) threshold determination using haze induced Chardonnay wine samples in a dark surround. The first stage was conducted over an elapsed period of eight days (six experimental days), the second stage was conducted over a single day. The experimental stages were separated by an elapsed period of around two months.

The information given below should be read in conjunction with the general methodology Section found in Chapter 3.

### 4.2.1 Preparation of microsphere suspensions

#### Achromatic microsphere suspensions

A primary microsphere suspension was created by mixing the 0.25 µm polymer microsphere concentrate (as supplied by the manufacturer) with HPLC grade deionised water. Two secondary base concentrates were then prepared in a similar fashion; these were further diluted to prepare a final set of 10 achromatic suspensions arranged in two-fold concentration steps. Turbidity details of the achromatic test samples are shown in Table 4.1.

**Table 4.1 Details of achromatic microsphere suspensions. Average turbidity values derived from repeated measurements taken over experimental period.**

| Achromatic samples | Average turbidity (NTU) |
|---|---|
| Diluent/Blank | 0.18 |
| 1 | 0.24 |
| 2 | 0.32 |
| 3 | 0.47 |
| 4 | 0.80 |
| 5 | 1.44 |
| 6 | 2.71 |
| 7 | 5.09 |
| 8 | 9.83 |
| 9 | 18.62 |
| 10 | 35.28 |

#### Yellow microsphere suspensions

Yellow microsphere suspensions were prepared in a similar fashion to the achromatic suspensions with the addition of yellow food colouring in the final stage diluent. Turbidity details of the yellow test samples are shown in in Table 4.2.

**Table 4.2 Details of yellow microsphere suspensions. Average turbidity values derived from repeated measurements taken over experimental period.**

| Yellow samples | Average turbidity (NTU) |
|---|---|
| Diluent/Blank | 0.18 |
| 1 | 0.24 |
| 2 | 0.32 |
| 3 | 0.49 |
| 4 | 0.77 |
| 5 | 1.39 |
| 6 | 2.58 |
| 7 | 4.99 |
| 8 | 9.68 |
| 9 | 18.38 |
| 10 | 34.44 |

### 4.2.2 Preparation of haze induced Chardonnay wine samples

The first part of this study had resulted in a haze detection threshold for yellow microsphere suspensions in a dark surround of 1.15 NTU (see Section 4.3.3). In an attempt to optimise the resolution of the wine threshold results, the range of wine sample turbidities (0.5 to 4.0 NTU) was designed to be much narrower than those used for the microsphere suspensions.

Wine stock exhibiting a high degree of visual haze was prepared by immersing 1 litre volumes of Chardonnay wine in a temperature controlled heat bath at 90˚C for a period of two hours; this formed a concentrate base for dilutions. Haze induced wine samples were prepared by diluting high haze wine stock with untreated wine using a 1.4 dilution factor; the sample turbidity outcomes are shown in Table 4.3.

**Table 4.3 Details of Chardonnay haze induced samples. Average turbidity values derived from measurements taken over experimental period**

| Chardonnay samples | Average turbidity (NTU) |
|---|---|
| Untreated/Blank | 0.4 |
| 1 | 0.60 |
| 2 | 0.79 |
| 3 | 0.89 |
| 4 | 1.10 |
| 5 | 1.25 |
| 6 | 1.54 |
| 7 | 2.06 |
| 8 | 2.39 |
| 9 | 2.98 |
| 10 | 3.57 |

### 4.2.3 Sensory booth configuration

In the first part of the study using microsphere suspensions the sensory booths were equipped and illuminated as previously described with two dark surround booths and two light surround booths being used.

For evaluation of the Chardonnay wine suspensions a single booth with a dark surround was used.

### 4.2.4 Microsphere suspension handling and presentation

For each booth a sample set consisting of 30 ISO glasses was prepared. This included 10 glasses each containing one of the 10 hazy samples and another 20 ISO glasses containing blank material; each glass contained 90 mL of liquid and was labelled with a unique random three digit code. The glasses were filled on the morning of the first day and refreshed with fresh samples on the mornings of the 3$^{rd}$ and 6$^{th}$ experimental days. Where stocks were not to be refreshed the next day, the glasses were stored overnight at room temperature with their tops covered to minimise evaporation.

Prior to each day's evaluations the exterior of the glasses were wiped clean using a lint free cloth.

### 4.2.5 Chardonnay wine sample handling, presentation and stability

The Chardonnay wine sample sets were prepared and presented in the same manner as the polymer microsphere suspensions except that all evaluations were completed in a single day thus avoiding the need to replenish samples.

Since the nominal sample range was well below 10 NTU, the possibility of excessive sample instability during subject evaluations was not anticipated (see Section 3.4.3); repeated turbidity measurements showed the samples to be stable over the experimental period.

### 4.2.6 Sensory evaluations of haze thresholds

#### Overview

Sensory evaluations of haze thresholds were conducted using the Ascending Method of Limits (AML) procedure[103] for both the microsphere suspensions and the wine samples. The 3-Alternate Forced Choice (3AFC) test[105] was used to regulate evaluation of each sample triplet.

#### Subjects

Subjects for the first part of the study consisted of 14 volunteers; 8 males and 6 females. For the second stage the subject pool consisted of 8 volunteers; 3 males and 5 females. All subjects involved in the second stage had participated in the first stage.

All subjects were staff, faculty members or post graduate students from Lincoln University, New Zealand who self-reported normal or corrected vision and were over the age of 18. None of the subjects had experience of the AML/3AFC procedure prior to the study. The subjects were not given any monetary or other incentive to participate in the study.

The subjects did not meet as a group; all evaluations were conducted individually.

#### Experimental design

In the first stage of the study there were four experimental booth configurations; see Table 4.4. There was one experimental configuration for the second stage involving wine samples: all wine evaluations were conducted in a single dark surround.

The presentation order of sample sets for each experimental configuration followed an ascending order of turbidity in accordance with the AML procedure. The order of the three glasses for each individual 3AFC evaluation was randomised with each possible order configuration used before re-use. In the case of the microsphere experiment, the order in which subjects were directed to individual booths followed a repeated 4x4 Latin Square.

**Table 4.4 Experimental configurations used for threshold determination in microsphere suspensions.**

| Configuration | Surround | Suspension colour |
|---|---|---|
| 1 | Dark | Achromatic |
| 2 | Dark | Yellow |
| 3 | Light | Achromatic |
| 4 | Light | Yellow |

**Evaluation procedure**

Prior to commencing evaluations each subject was given a formal pre-evaluation briefing that lasted approximately 5 minutes. The briefing explained the AML/3AFC procedure and was read from pre-written scripts (Appendix C.1 and C.2). The subject was told that they would be repeatedly presented with sets of three glasses and for each set they were to determine which glass held the haziest sample. The subject was told they were to evaluate the samples by first viewing each of the three possible, two-glass combinations at eye level after which they were to be allowed to manipulate the glasses freely (i.e. tipping the glass was permitted). Towards the end of the briefing the subject was requested to identify a mid-range hazy sample from a set of three glasses that contained two blanks; all subjects successfully identified the hazy sample. The subject was then directed to the appropriate sensory booth to start the evaluation process.

Sets of three glasses were presented to the subjects for each 3AFC evaluation; one contained a test sample whilst the others contained blank material. Ten sets of three glasses were presented in each experimental configuration (i.e. ten 3AFC tests per configuration). The subjects were given 30 seconds to make their evaluations at which point the 3AFC triplet was retrieved and replaced by the next triplet. Once all ten 3AFC evaluations had been completed for a particular experimental configuration, the subjects were directed to the next booth. The process continued until all samples in every booth had been evaluated once. In the case of the wine samples, one set of 10 evaluations in the single dark booth completed the evaluation session for the subjects.

Each experimental booth contained its own combined instruction/score sheet that remained visible during the evaluations in that booth.

## 4.3 Results and discussion

### 4.3.1 Threshold quantification

Individual best estimate thresholds (BET's) were calculated as the geometric mean of turbidities between an incorrect response and a correct response that was only followed by correct responses. The group threshold BET for a particular experimental configuration was calculated as an overall turbidity value by taking the geometric mean of all the individual BET's for that configuration.

### 4.3.2 Raw data assessment

The individual BET data for the microsphere evaluations was found not to be normal with each configuration showing a mix of positive skew and bi-modal attributes. The distribution for yellow microsphere suspensions in the dark surround is shown in Figure 4.1; this is representative of the

data for the other three microsphere configurations. The BET data for the haze induced Chardonnay wine samples was found to take on a more normal form when plotted with the same bin steps used in the microsphere histogram; this is shown in Figure 4.2. Narrowing the bin steps in the wine histogram leads to a similar bi-modal form as seen for the microsphere distribution. There are insufficient BET scores to be confident of the nature of the distributions.

In the single previous study where data distributions were reported[54] a mix of bi-modal and positively skewed data was also found. In general, previous studies[3,5,6] have log transformed the data prior to applying statistical analysis without disclosing the resulting distributions. An attempt to normalise the microsphere data in this study through log transformation was not successful.

### 4.3.3 Observed thresholds

The dark surround and achromatic suspension configuration resulted in the lowest group threshold: 0.44 NTU. The effect of changing to a light surround appeared sizeable with the group threshold increasing to 3.24 NTU. The yellow suspensions followed a similar pattern with thresholds increasing from 1.15 to 6.93 NTU as the surround was changed from dark to light. A group threshold of 0.86 NTU was found for the Chardonnay wine samples; this is somewhat smaller than the equivalent yellow microsphere threshold of 1.15 NTU. The group BET's and individual subject range for each of the experimental configurations are detailed in Table 4.5.

**Table 4.5 Group Best Estimate Threshold (BET) and individual subject range for each experimental configuration.**

| Sample material | Surround | Sample colour | Group BET (NTU) | Individual ranges (NTU) |
|---|---|---|---|---|
| Polymer microspheres | Dark | Achromatic | 0.44 | 0.21 to 1.07 |
| Polymer microspheres | Dark | Yellow | 1.15 | 0.4 to 3.59 |
| Polymer microspheres | Light | Achromatic | 3.24 | 1.07 to 13.53 |
| Polymer microspheres | Light | Yellow | 6.93 | 1.89 to 25.16 |
| Chardonnay wine | Dark | Yellow (natural) | 0.86 | 0.52 to 1.39 |

It can be seen that the absolute ranges in the subject responses expand substantially as the group BET rises. However, the individual ranges were consistently spread over six concentration levels for each of the five experimental configurations. The Chardonnay wine samples were prepared using a smaller dilution factor than the polymer microsphere samples (1.4 versus 2) meaning that the relative extent of the individual range in the wine samples was somewhat narrower than that observed in the equivalent microsphere samples.

Figure 4.3 shows the results of this study for achromatic samples in comparison to those found by Fleet and Siebert[6] who used a brightly illuminated viewing box with similar particle sizes. Although

differential significance cannot be objectively judged, the more realistic viewing conditions of this study appear to lead to lower thresholds in the light surround. Superficially the reverse appears to be the case for the dark surrounds but the differences are less pronounced. However, caution is required when comparing absolute NTU values as inconsistencies in inter-device nephelometric measurements may be misleading (see Section 2.2.2). Figure 4.4 shows the most comparable results available for coloured suspensions (Fleet and Siebert[6]) where a similar pattern emerges to that seen for the achromatic suspensions.

### 4.3.4 Non parametric statistics

A series of Wilcoxon Signed Rank tests[115] were used to determine whether there were significant differences in medians between the data for the two types of surround and the two types of suspension (Table 4.6). It can be seen that changing the suspension colour in the black surround did not result in a significant change in median. In contrast, the same suspension change in the light surround was found to be highly significant. The substantial impact of the surround is evidenced by the significance of the difference between the dark and light surrounds for both suspension types.

It may be of some interest that the only test where surround induction was not expected to have a had a substantial influence in at least one of the configurations was the one that shows no significant difference in the median. In other words the lack of significance between the two configurations in Test 1 may be related to the fact that the surround luminance was less than all test sample luminances in both configurations (see Section 2.4.2 for more on induction effects).

**Table 4.6 Results of Wilcoxon Signed Rank test for median changes in surrounds and suspension types.**

| Test | Experimental configuration A | Experimental configuration B | Significant difference in median ($\alpha$=0.01)[1] |
|---|---|---|---|
| 1 | Dark surround, achromatic suspension | Dark surround, yellow suspension | No |
| 2 | Light surround, achromatic suspension | Light surround, yellow suspension | Yes |
| 3 | Dark surround, achromatic suspension | Light surround, achromatic suspension | Yes |
| 4 | Dark surround, yellow suspension | Light surround, yellow suspension | Yes |

[1]None of the relevant datasets included more than 25% ties.

A Wilcoxon Signed Rank test was also used to explore the possible differences between the Chardonnay wine samples and yellow polymer microsphere suspensions viewed in the dark surround. It was possible to conduct this test because all eight subjects involved in the wine evaluations also took part in the yellow polymer microsphere evaluations under equivalent

experimental conditions. It is acknowledged, however, that the experiments were conducted some two months apart and this may have undermined the pairing of subjects' results. An equal number of positive and negative differences were found across the eight subjects; no significant difference was found.

### 4.3.5 Summary and conclusions

It appears that both pigment content and surround viewing conditions influence haze detection thresholds. These results generally support those of Fleet and Siebert[6] whose data showed that thresholds substantially increase as the surround becomes lighter. The impact of suspension colour in other studies has not been consistent but these results suggest yellow pigments will tend to cause an increase in haze detection thresholds in lighter surrounds.

In terms of absolute threshold levels under real-world conditions, Horne et al.[4] reported a threshold of around 3.55 NTU for red/yellow formazin suspensions. Unfortunately, there is no information regarding the surround viewing conditions employed making any comparison with this study's results difficult. The best that may be said is that the results found in this study for light and dark surrounds are perhaps consistent with viewing extremes and that Horne et al. may have employed some intermediate conditions.

The lower BET for the Chardonnay wine samples along with the narrower subject range compared to the equivalent yellow microsphere samples suggests that more effective threshold cues may be present in the wine samples used. The apparent difference in the results may be associated with chromatic differences or visual artefacts caused by the different compositions. However, the Wilcoxon Signed Rank test did not detect any significant difference between the results and more work is required to investigate this further.

In general, a dark surround should be considered as one in which haze detection thresholds are likely to be the lowest for any given yellow sample; a threshold of around 1.0 NTU would seem a reasonable assumption under such conditions. This conclusion is specific to the lighting, samples and receptacles used in this study. Generalisation of these results to other viewing conditions and sample compositions cannot be made without further work.

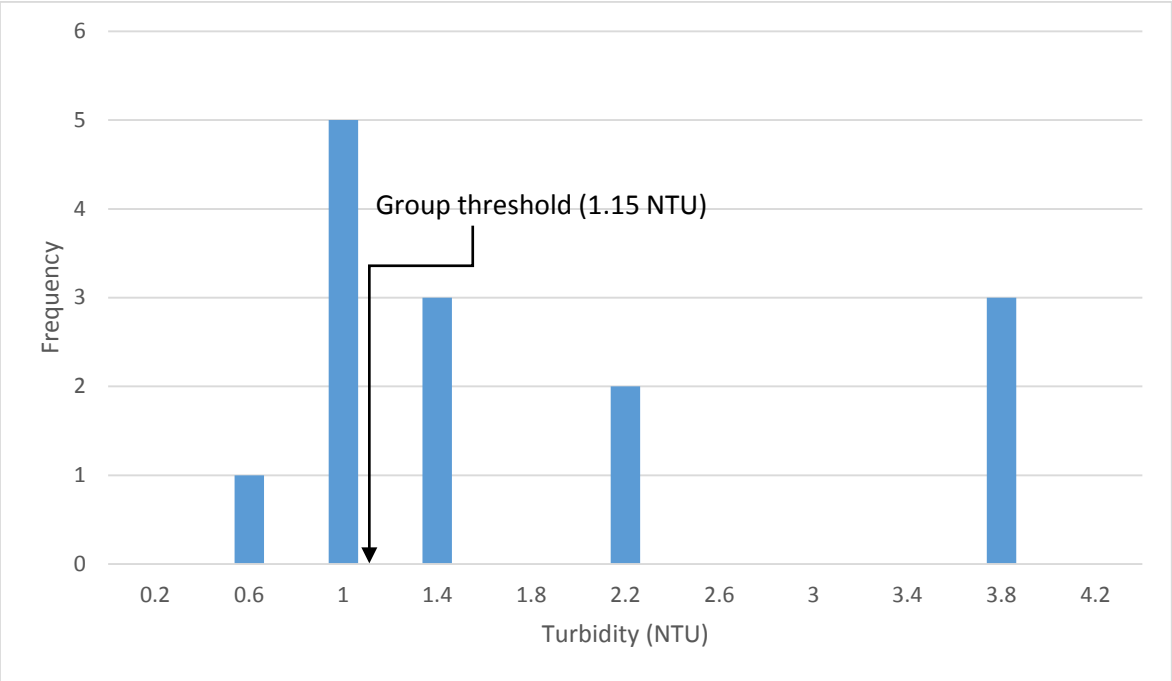


**Figure 4.1** **Histogram of individual haze detection BET thresholds for yellow microsphere suspensions evaluated in a dark surround (n=14).**

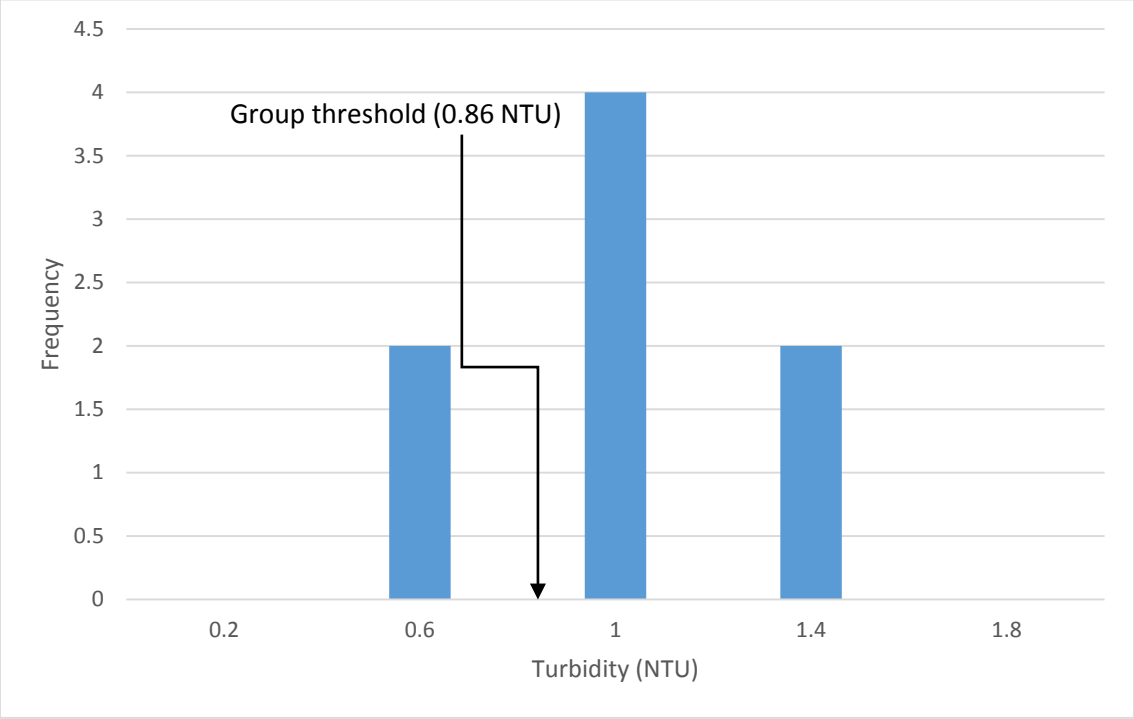


**Figure 4.2** **Histogram for individual haze detection BET thresholds in haze induced Chardonnay wine samples evaluated in a dark surround (n=8).**

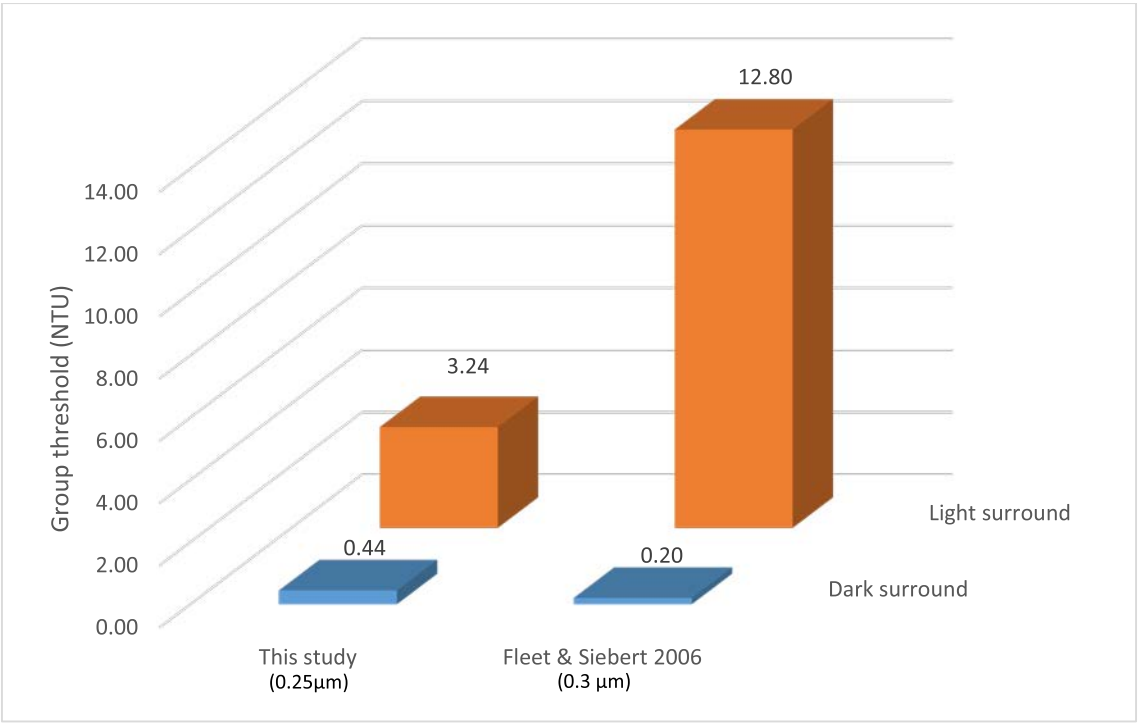


**Figure 4.3** **Inter-study comparison of haze detection thresholds in achromatic microsphere suspensions viewed in dark and light surrounds. Data from Fleet and Siebert[6] derived under closely controlled viewing conditions using similar AML/3AFC tests.**

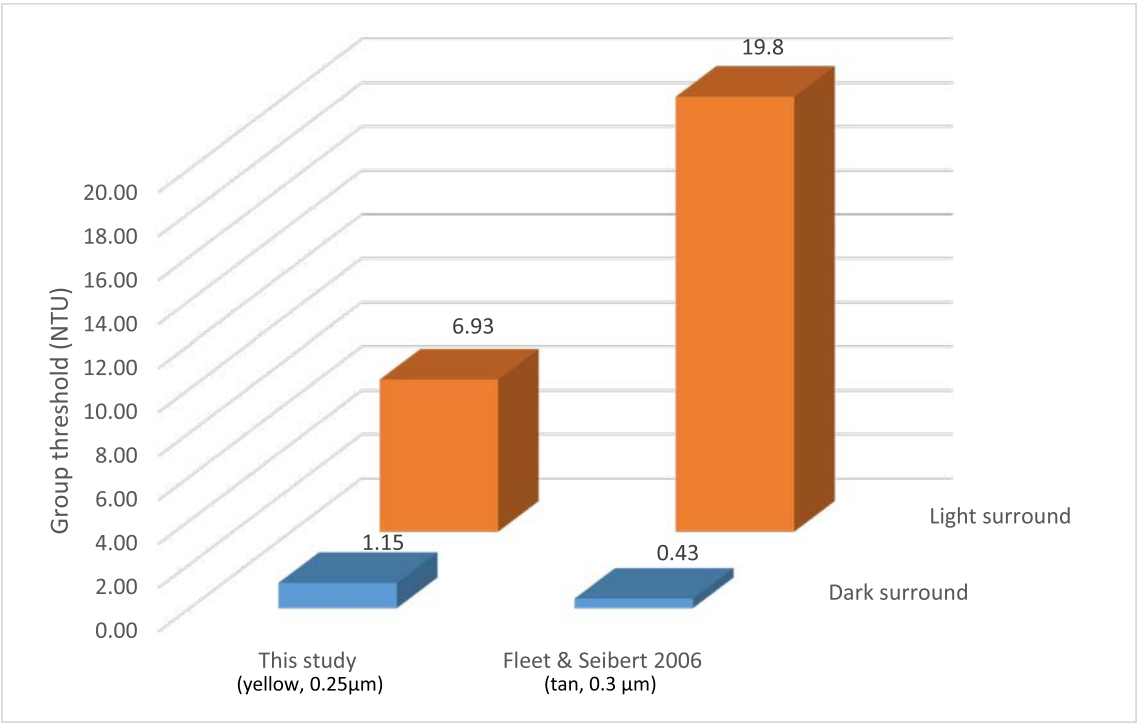


**Figure 4.4** **Inter-study comparison of haze detection thresholds in coloured samples viewed in dark and light surrounds. Data from Fleet and Siebert[6] derived under closely controlled viewing conditions using similar AML/3AFC tests. Comparable data for yellow suspensions not available; tan coloured suspensions offer the closest match.**

# Chapter 5

# Relationships between Haze Intensity Perception, Nephelometry and Aggregate Luminance in Polymer Microsphere Solutions

## 5.1 Overview and objectives

The overall goal of this study was to investigate the human response to haze intensity and to examine the relationship between perceptual and instrumental data for samples containing polymer microspheres of a single nominal diameter (0.25 μm). The underlying hypothesis was that human perception of haze correlates with the common use terms of *brightness* and *darkness* and that one of the many models proposed for scaling perceptual brightness or lightness may offer a means of quantifying perceptual haze intensity (see Section 2.3.3).

The primary objective was to determine the haze intensity response of achromatic and yellow samples containing polymer microspheres using a panel of human subjects and to compare that response to those predicted by the scaling models outlined previously (Sections 2.5 and 3.5.2). In contrast to most related work, the viewing conditions were to be designed to match those of real-world scenes as far as possible whilst controlling key variables.

A secondary objective was to investigate the relationship between microsphere concentrations, instrumental turbidity and image luminance. Of particular interest was the possibility that digital image data might offer a means of tracking instrumental turbidity.

It was decided to use polymer microspheres rather than wine samples since this provided better control over experimental conditions and sample. This study can be viewed as a necessary step prior to engaging in similar work using *bona fide* wine samples (see Chapter 6).

## 5.2 Materials and methods

The information given below should be read in conjunction with the general methodology section found in Chapter 3.

### 5.2.1 Sample preparation

The results of the threshold detection experiment (Chapter 4) had shown that thresholds for the yellow suspensions were higher than those for achromatic suspensions. Further pilot work was conducted and separate two-fold dilution schemes were designed for the yellow and achromatic suspensions with the lowest step in the range just above the previously observed thresholds.

### Achromatic microsphere suspensions

A primary microsphere suspension was created by mixing 2.0 mL of the polymer microsphere solution (as supplied by the manufacturer) with 50 mL of HPLC grade deionised water and sodium dodecyl sulphate (SDS) at 0.1%. The resulting suspension was then placed on a rocker mechanism and gently rocked for 3 hours to ensure microsphere dispersal and to aid surfactant efficacy. The primary microsphere solution was then dialysed (see Section 3.4.4) resulting in an achromatic microsphere suspension with a dry mass microsphere concentration of 0.04 g/mL and an SDS concentration of 0.1%. Subsequently two further secondary base suspensions were then prepared at concentrations of $1.69 \times 10^{-3}$ g/mL and $5.27 \times 10^{-5}$ g/mL with an SDS concentration of 0.005%. The two secondary base suspensions were then used to prepare a final set of 10 achromatic suspensions arranged in two-fold concentration steps with an SDS concentration of 0.005%; details in Table 5.1

**Table 5.1 Details of achromatic sample suspensions. Average turbidity values derived from measurements taken over experimental period.**

| Achromatic sample | Microsphere density (number per mL) | Dry mass concentration (g/mL) | Average turbidity (NTU) |
|---|---|---|---|
| 1 | $1.54 \times 10^{7}$ | $1.32 \times 10^{-7}$ | 1.1 |
| 2 | $3.07 \times 10^{7}$ | $2.64 \times 10^{-7}$ | 2.1 |
| 3 | $6.14 \times 10^{7}$ | $5.27 \times 10^{-7}$ | 4.1 |
| 4 | $1.23 \times 10^{8}$ | $1.05 \times 10^{-6}$ | 7.8 |
| 5 | $2.46 \times 10^{8}$ | $2.11 \times 10^{-6}$ | 15.0 |
| 6 | $4.91 \times 10^{8}$ | $4.22 \times 10^{-6}$ | 28.1 |
| 7 | $9.82 \times 10^{8}$ | $8.44 \times 10^{-6}$ | 52.7 |
| 8 | $1.96 \times 10^{9}$ | $1.69 \times 10^{-5}$ | 92.9 |
| 9 | $3.93 \times 10^{9}$ | $3.38 \times 10^{-5}$ | 158.5 |
| 10 | $7.86 \times 10^{9}$ | $6.75 \times 10^{-5}$ | 251.3 |

### Yellow microsphere suspensions

The yellow microsphere suspensions were prepared in a similar fashion to the achromatic suspensions with the addition of yellow food colouring (0.25%) in the diluents. The outcome was a final set of 10 suspensions arranged in two-fold concentration steps with an SDS concentration of 0.005%.

**Table 5.2 Details of yellow sample suspensions. Average turbidity values derived from measurements taken over experimental period.**

| Yellow sample | Microsphere density (number per mL) | Dry mass concentration (g/mL) | Average NTU |
|---|---|---|---|
| 1 | $2.40x10^{7}$ | $2.06x10^{-7}$ | 1.6 |
| 2 | $4.80x10^{7}$ | $4.12x10^{-7}$ | 3.1 |
| 3 | $9.59x10^{7}$ | $8.24x10^{-7}$ | 6.0 |
| 4 | $1.92x10^{8}$ | $1.65x10^{-6}$ | 11.6 |
| 5 | $3.84x10^{8}$ | $3.30x10^{-6}$ | 22.2 |
| 6 | $7.68x10^{8}$ | $6.59x10^{-6}$ | 42.3 |
| 7 | $1.54x10^{9}$ | $1.32x10^{-5}$ | 76.0 |
| 8 | $3.07x10^{9}$ | $2.64x10^{-5}$ | 131.5 |
| 9 | $6.14x10^{9}$ | $5.27x10^{-5}$ | 213.0 |
| 10 | $1.23x10^{10}$ | $1.05x10^{-4}$ | 330.0 |

### 5.2.2 Sensory booth configuration

Sensory booths were equipped and illuminated as previously described with two dark surround booths and two light surround booths being employed in this study.

### 5.2.3 Allocation of sample ranges and suspension types to booths

A prior study (Chapter 4) had shown that threshold detection levels for yellow suspensions were higher than those for achromatic suspensions and that thresholds were also higher in light surrounds compared to those in dark surrounds. The allocation of concentration ranges to experimental configurations was designed to commence at a level just around the best estimate group thresholds previously determined and to extend in two-fold geometric steps. The relevant range and modulus concentration values are shown in Table 5.3.

**Table 5.3 Allocation of stocks to experimental booth configurations. Modulus refers to the reference sample used in Magnitude Estimation.**

| Booth Surround | Suspension Colour | Lowest Concentration (spheres/mL) | Highest Concentration (spheres/mL) | Modulus Concentration (spheres/mL) |
|---|---|---|---|---|
| Dark | Achromatic | $1.54 x10^{7}$ | $1.96 x10^{9}$ | $2.46 x10^{8}$ |
| Dark | Yellow | $2.40 x10^{7}$ | $3.07 x10^{9}$ | $3.84 x10^{8}$ |
| Light | Achromatic | $6.14 x10^{7}$ | $7.86 x10^{9}$ | $9.82 x10^{8}$ |
| Light | Yellow | $9.59 x10^{7}$ | $1.23 x10^{10}$ | $1.54 x10^{9}$ |

### 5.2.4 Sample handling and presentation

A sample set consisting of eight ISO glasses each containing 90 mL of microsphere suspension was prepared for use in each booth along with a mid-range reference modulus containing the same

volume; each glass was labelled with a unique random three digit code. These were prepared on the morning of the first day and were refreshed with fresh stock on the morning of the 3rd day. The glasses were stored overnight at room temperature with their tops covered to minimise evaporation. Prior to each day's evaluations the exterior of the glasses were wiped clean using a lint free cloth.

### 5.2.5 Sensory evaluations of haze intensity

#### Overview

Sensory evaluations of haze intensity were conducted using the technique of Magnitude Estimation[105,116] using a continually present reference modulus. The evaluation sessions were scheduled for a duration of one hour and were spread over an elapsed period of four days.

#### Subjects

Subjects for this study consisted of 15 volunteers (staff, faculty members or post graduate students) from Lincoln University, New Zealand. All subjects self-reported normal or corrected vision. The subject pool included 5 males and 10 females all over the age of 18. None of the subjects had experience of the Magnitude Estimation technique prior to the study. The subjects were not given any monetary or other incentive to participate in the study.

#### Experimental design

The presentation order of samples followed a repeated 8x8 Latin Square structure so that each of the eight samples appeared in each initial position and in every other position an equal number of times before re-using the presentation order. The repeated, 8x8 Latin Square design was re-used across the four booths by rolling the order so that the subjects encountered a different order in each booth. The order in which subjects were directed to individual booths followed a repeated 4x4 Latin Square. The hand with which the modulus was held during evaluations was alternated by subject. Each subject evaluated every sample in every booth configuration once.

#### Evaluation sessions

Prior to commencing sample evaluations each subject was trained individually in the Magnitude Estimation technique using shapes and areas in the manner recommended by Moskowitz[116]. The subject was then told that the attribute to be evaluated was the "level of haze" in a liquid test sample. The subject was given the opportunity to perform trial evaluations on samples similar to those to be encountered in the formal evaluations. All subjects successfully demonstrated an understanding of the task by suitably rating trial samples against a modulus reference.

Once the training session was complete, the subject was then directed to the sensory booth area and shown a combined instruction/score sheet along with a formal pre-evaluation briefing that lasted approximately five minutes. The briefing was read from a pre-written script (Appendix C.3) and

served to reinforce the training and to also allow the subject's vision to adapt to the sensory booth lighting. During the briefing the subject was shown a sample and told it was to act as a reference with an intensity value of 100 against which all test samples were to be judged (samples with one tenth the haze of the reference were to be assigned the value of 10, samples with twice the haze of the reference were to be assigned the value of 200). Zero evaluations were allowed and no upper limit was imposed. Each sample was then presented to the subject in the order and manner previously described and the subject recorded evaluations on a score sheet common to all samples in that booth. The subjects were given 30 seconds to make their evaluations at which point the sample was retrieved and replaced by the next sample. The modulus remained present in the booth at all times. Once all eight samples for a particular booth had been evaluated, the subject was directed to the next booth. The process continued until all samples in every booth had been evaluated once.

**Perceptual quantification and zero handling**

The intensity of haze perception was quantified by taking the geometric mean of the magnitude estimates given by the subjects for each concentration/booth surround/suspension colour combination resulting in eight geometric mean values for each of the four surround/colour combinations. Concentration levels were excluded from the final results where the number of zero values assigned exceeded 50% of the number of subject evaluations for that level. Concentration levels with less than or equal to 50% zero evaluations had the zeroes replaced with half the lowest, non-zero value for that level to allow the geometric mean to be calculated.

### 5.2.6 Nephelometric measurements

Nephelometric measurements were conducted on the sample stocks in the manner previously described. Measurements were made on each day of the study providing five turbidity values for each the 20 stocks over the experimental period.

### 5.2.7 Image capture

On the final day of the study digital images were captured using a modified Canon A2300 camera and processed in the manner previously described.

**Scene arrangement**

Two of the booths used for sensory evaluations (one with a dark surround and one with a light surround) were modified to allow images to be captured of samples in a scene that closely resembled that encountered by the human subjects. This involved the installation of a dark linen covered platform (28 cm x 13 cm) at a height approximating the subjects' eye level (510 mm). Each image scene consisted of 2 ISO glasses placed centrally upon the dark linen platform separated by a

distance of 60 mm. Pin guides were placed on the platform to ensure consistency of glass placement between images. Booth surrounds and illumination were otherwise left unchanged.

**Imaging sequence**

In each image scene the contents of the glasses consisted of a suitable modulus reference and a sample to be assessed. The modulus was included to emulate inter-glass illumination that may have been present during the sensory evaluations. Each of the 10 sample suspensions for each colour (10 achromatic, 10 yellow) were imaged twice in each booth surround (modulus and sample positions being swapped for each sample concentration). The image sequence was randomised for each suspension colour and booth type. This resulted in 80 images (4 randomised sets of 20 images for each colour/surround combination).

**Image capture and processing**

The camera was positioned orthogonally at 490 mm from the centre line of the glasses and a zoom setting established that ensured that the samples were positioned within the central 50% of the image view to minimise vignetting effects. An exposure time of 31 mS was used for all images; all other camera parameters were as described in A.3 . Processing of the images to arrive at aggregate luminance assessments sample were conducted in the manner previously described with a Region of Interest (ROI) consisting of 400x400 pixels located in an upper central position on the glass. The final aggregate luminance value for each sample was derived by arithmetically averaging the two values captured for the left and right positions.

### 5.2.8 Data analysis and curve fitting

Non-linear regression techniques were used to fit the experimental data to various generalised brightness/lightness models as previously described (Section 3.5.2). Aggregate luminance was used as the independent variable and the geometric means of subject evaluations as the dependent variable. Each surround/suspension type combination was analysed individually in this manner.

## 5.3 Results and discussion - perceptual haze intensity and aggregate luminance

### 5.3.1 Review of subjects' evaluation data

**Zero evaluations and impact on datasets**

The decision to extend the concentration levels close to the known detection thresholds was expected to present some difficulties for the subjects and this was reflected in the number of zero evaluations assigned; the distribution of zero evaluations across the booth configurations is shown Figure 5.1.

The dark surround / achromatic suspension booth was the only configuration where the lowest sample concentration was higher than the *highest* individual subject threshold previously observed. It might reasonably have been expected that the zero count in dark/achromatic booth would have been the lowest of the configurations since 73% of the subjects in this study had taken part in the previous threshold study. The fact that the black surround/achromatic suspension experienced the highest zero count may indicate that the threshold detection mechanisms at work in the lowest absolute ranges are not of a nature that supports perceptual scaling. Visual cues such as small surface highlights may provide information regarding the presence of haze without offering an adequate visual area to allow scaling.

The zero evaluations were processed as previously described and this resulted in a reduction in the number of useable concentration levels in the overall dataset (details are provided in Table 5.4).

**Table 5.4 Impact of zero evaluations on useable concentration levels.**

| Booth Surround | Suspension Colour | Nominal number of levels | Useable number of levels |
|---|---|---|---|
| Dark | Achromatic | 8 | 6 |
| Dark | Yellow | 8 | 7 |
| Light | Achromatic | 8 | 7 |
| Light | Yellow | 8 | 7 |

### Discrimination errors

In Magnitude Estimation studies the steps between stimuli should be chosen such that each step is easily discernable by the subjects. However, it is possible that part of a range may be more easily distinguishable than another and examination of discrimination errors can be used to identify problematic areas in the evaluation data. A discrimination error is observed when a subject gives a rating to higher concentration that is less than or equal to a rating assigned for a lower concentration. Discrimination errors for this study are detailed in Table 5.5.

**Table 5.5 Mean discrimination errors for each experimental configuration.**

| Booth Surround | Suspension Colour | Mean discrimination errors (%) |
|---|---|---|
| Dark | Achromatic | 1.1 |
| Dark | Yellow | 4.8 |
| Light | Achromatic | 2.9 |
| Light | Yellow | 3.8 |

Error rates for comparable vision studies have not been found but the rates in this study are substantially lower than those reported for taste studies[117,118] where error levels of 7% are not uncommon. Individual errors were examined for evidence of clustering at particular concentration

levels. No clustering was apparent with no more than two errors at any one level. However, the majority of these errors occurred at the higher concentration levels suggesting a possible saturation of response at those levels.

### Inter-subject variation

The inter-subject variation was calculated using the Coefficient of Variance (CV) method previously described using the geometric mean as the base reference. These calculations generated a CV value for each subject for each booth configuration. The overall arithmetic mean of the subjects' CV values are listed in Table 5.6. Although the CV differences between the configurations are small, it appears that yellow pigment in the samples may result in elevated error rates.

**Table 5.6 Inter-subject variation by booth configuration**

| Booth Surround | Suspension Colour | CV Mean (%) |
|---|---|---|
| Dark | Achromatic | 15.4 |
| Dark | Yellow | 22.4 |
| Light | Achromatic | 18.6 |
| Light | Yellow | 21.7 |

Other studies involving unrelated colours report a range of CV mean values for brightness evaluation by Magnitude Estimation as ranging from 11% [59] to 48%[106]. A study of related colours by Martin et al. found the mean of CV values at 17% for lightness[34]. Differences in experimental conditions prevent direct comparisons but the study by Martin et al. is the closest comparable study for the light booth surrounds and Luo and Hunt[119] state that a CV value of 13% for lightness and 10% for brightness are typical in related colour studies. The reason for the slightly elevated CV values is unclear but may be related to chromatic effects since CV values increase with chromatic content (light surrounds exhibit a bias in the green channel, yellow suspensions reduce blue channel values).

In order to investigate variances for bias or unusual ratings, it is useful to examine CV values at an individual subject level. These are best viewed graphically and can be seen in Figure 5.2. It is clear that Subject 3 is the source of some unusual variance with extreme CV values for the light booth configurations. Comparisons between Subject 3 evaluations, Subject 11 evaluations (the subject with least variance) and the geometric mean are shown in Figure 5.3. In the lower part of the ranges Subject 3's evaluations are broadly in accordance with other subjects. In the upper part of the ranges it appears that Subject 3's basis for evaluation deviates from that employed by the other subjects. It is possible that Subject 3 could not construct appropriate ratio evaluations at the higher ranges and that the ratings given are just markers of extremes[120]. For the purposes of analysis it was assumed that Subject 3's data was an example of non-ratio scoring in the higher ranges. Subject 3's data has therefore been excluded in the analysis below.

### 5.3.2 Dark surround and achromatic suspensions

The relationship between perceived haze and aggregate luminance for the dark surround and achromatic suspensions was found to exhibit a response consistent with a smooth compressive function. Non-linear regression was applied to the data and a good fit was found using a power function; a marginally lesser fit was provided by a log function. Regression results are summarised in Table 5.7; data and best fit curves shown in Figure 5.4. The value of the exponent (0.45) for the power function is generally consistent with adaptation to illuminated conditions [93,121]. It is also consistent with the data from Carrasco[54] linking haze perception to instrumental turbidity for comparable turbidity ranges (see Section 2.3.2). Although the subject variation for the data was low at a CV of 15.4, the validity of the 0.45 power exponent is somewhat undermined by the absence of data for the lowest levels that were thought to be above threshold detection.

**Table 5.7 Results for non-linear regression of haze perception versus aggregate luminance data for dark surround and achromatic suspensions (aggregate luminance values in normalised form). Log and power models with highest $R^2$ values shown.**

| Model | Form | k | b | d | n | $R^2$ |
|---|---|---|---|---|---|---|
| Eqn 3.3 | $y = k(x-b)^n$ | 35.9 | 19.5 | NA | 0.45 | 0.9991 |
| Eqn 3.7 | $y = k\log(1 + d(x-b)/b)$ | 200.4 | 18.6 | 3.54 | NA | 0.9977 |

The dark surround was not expected to cause any gross induction effects since its luminance level was below that of the lowest luminance sample but previous work has shown that the increment/decrement border is not exactly aligned to the luminance of the surround[93]. To investigate whether surround induction had been active, the data was transformed to a log-log basis (as is common in these types of studies) where a power function with a constant exponent should be linear. It can be seen in Figure 5.5 that there is sharp increase in the slope commencing somewhere below the level of the modulus. Although the log transformation works to amplify any small deviance from the power response, this may be evidence that the presence of the modulus whilst subjects were evaluating samples had an inductive effect on the results. Alternatively, it may simply reflect the fact that the true response is not a simple power function close to the threshold.

### 5.3.3 Dark surround and yellow suspensions

The relationship between perceived haze and aggregate luminance values for the dark surround and yellow suspension was found to exhibit a broadly similar response to the achromatic suspensions in the dark surround. Non-linear regression results are summarised in Table 5.8; data and best fit curves shown in Figure 5.6. Although the $R^2$ value for Equation 3.3 appears relatively high, the fit deviates systematically from the data; the log function of Equation 3.7 is the superior model. Again the log-log

plot of perceived haze versus aggregate luminance shows a possible bi-segment response with the modulus luminance located at the shift in slope (Figure 5.7).

**Table 5.8 Results for non-linear regression of haze estimation versus aggregate luminance data for dark surround and yellow suspensions (aggregate luminance values in normalised form). Log and power models with highest $R^2$ values shown.**

| Model | Form | k | b | d | n | $R^2$ |
|---|---|---|---|---|---|---|
| Eqn 3.7 | $y = k\log(1 + d(x-b)/b)$ | 248.5 | 13.8 | 1.74 | NA | 0.9988 |
| Eqn 3.3 | $y = k(x-b)^n$ | 26.6 | 14.0 | NA | 0.53 | 0.9919 |

### 5.3.4 Light surround and achromatic suspensions

The relationship between perceived haze and aggregate luminance for the light surround and achromatic suspensions was found to exhibit a broadly similar response to that found in the equivalent dark surround. It is possible that an expansive lower segment is followed by a mildly compressive upper segment with a transition occurring somewhere in the region of the luminance of the modulus (Figure 5.8).

Non-linear regression was applied to the entire data range and the results are summarised in Table 5.9; data and best fit curves shown in Figure 5.9. Although both power and log models display a high $R^2$ value, neither model appears to reflect the true structure of the data. Figure 5.10 shows the data in log-log form; the upper segment is close to linear and appears to conform to a power function (with an exponent of approximately 1.2) but the lower segment does not.

**Table 5.9 Results for non-linear regression of haze estimation versus aggregate luminance data for light surround and achromatic suspensions (aggregate luminance values in normalised form). Log and power models with highest $R^2$ values shown.**

| Model | Form | k | b | d | n | $R^2$ |
|---|---|---|---|---|---|---|
| Eqn 3.7 | $y = k\log(1 + d(x-b)/b)$ | 320.3 | 34.8 | 3.0 | NA | 0.9972 |
| Eqn 3.3 | $y = k(x-b)^n$ | 21.3 | 34.9 | NA | 0.61 | 0.9930 |

### 5.3.5 Light surround and yellow suspensions

The relationship between perceived haze and aggregate luminance for the light surround and yellow suspensions is shown in Figure 5.11. In contrast to the achromatic suspensions in the same surround, both the lower and upper segments appear somewhat expansive. The segment transition remains around the level of the modulus.

Non-linear regression was applied to the entire data range and the results are summarised in Table 5.10; data and best fit curves in Figure 5.12. Although high $R^2$ value fits can be achieved using both

log and power functions, they do not appear to fit the structure of the data. Figure 5.13 shows the data in log-log form. The upper segment is strongly linear and appears to conform to a power function (with an exponent of approximately 1.3). Again the lower segment is not linear and does not correspond to a power function.

The presence of the yellow pigment appears to accelerate the response in both the lower and upper segments; this would be consistent with the Helmholtz–Kohlrausch effect[75] whereby more colourful stimuli are perceived as lighter for the equivalent level of sample luminance.

**Table 5.10 Results for non-linear regression of haze estimation versus aggregate luminance data for light surround and yellow suspensions (aggregate luminance values in normalised form). Log and power models with highest $R^2$ values shown.**

| Model | Form | k | b | d | n | $R^2$ |
|---|---|---|---|---|---|---|
| Eqn 3.3 | $y = k(x-b)^n$ | 10.9 | 31.0 | NA | 0.76 | 0.9944 |
| Eqn 3.7 | $y = k\log(1 + d(x-b)/b)$ | 830.1 | 29.8 | 0.49 | NA | 0.9926 |

### 5.3.6 Summary and conclusions

In the absence of a generally accepted scaling model for perceptual brightness/lightness in real-world conditions, it has been necessary to adopt non-linear regression and curve fitting as an investigative method. Although model variants have been found that appear to fit the data well, generalisation of the models to useful forms requires a means of deriving the various parameters involved. In the case of Stevens Power Law it is common practice to cite the power function exponent as some form of single, critical value for brightness perception whilst ignoring the other parameters. Stevens and Marks[122] emphasise the dependence of the response on all the parameters as determined by the specific viewing conditions and adaptive state of the visual system. Practical application of a power function model to estimate perceptual haze would require knowledge of the exponent, intercept and threshold. There is no simple way to estimate these parameters for complex real-world scenes. Similar comments may be made for the various logarithmic models such as Log W since the generalised log fits did not exhibit constant parametric values over the different experimental configurations.

The general hypothesis that haze perception might be quantified through the use of a simple power or logarithmic function seems not to be acceptable even if the relevant model parameters were knowable. This is because any real scene will be composed of multiple elements any of which may influence the perception of the target sample[74,123]. In each of the experimental configurations used in this study it is probable that such an effect has biased the observed data when the sample luminance has been less than that of the modulus luminance. The consistent deviance of the response at or around the level of the modulus suggests that self-illuminating elements positioned in the same

plane as the target sample may have a powerful inductive effect on adjacent elements. It is difficult to distinguish the relative inductive effects of the surround from that of the modulus, but it is possible that the light surround has caused a steepening of the data around the luminance of the modulus in Figure 5.9 and Figure 5.12. This may be evidence of the Crispening Effect reported by other authors[68,70].

An additional issue relates to the compound nature of haze perception. At high levels of haze some correspondence with achromatic brightness perception may be expected but at low levels this assumption seems less plausible. As the haze in any real sample approaches threshold levels, it is likely that transparency and particulate detection will become increasingly important and these will likely scale separately from perceptual brightness or lightness. It is also unclear how chromatic effects will be construed with respect to haze but the manner in which subject variance appeared to increase with colour content suggests that it will have some material impact.

The general models used in this study were elaborated from relatively simple empirical models of brightness and lightness developed under achromatic conditions. It is clear they are not sufficiently sophisticated to cope with the complexity of coloured samples being viewed in real-world scenes composed of multiple elements. Progress in this area is likely to require the use of a full colorimetric appearance model specifically designed to accommodate inductive effects; CIECAM-m2[101] is one such candidate.

Any future work should be designed carefully such that experimental interferences are minimised. Specifically, modulus reference materials should not be present during test sample evaluation unless those materials are intended to be experimental variables. Care should also be given to the glassware that subjects may use to correct their vision; tinted or activated lenses should be avoided since they are likely to add to experimental variance.

## 5.4 Results and discussion - microsphere concentrations, turbidity and aggregate luminance

### 5.4.1 Microsphere concentration and turbidity

The relationship between microsphere concentration and instrumental turbidity was found to follow a smooth, largely compressive function for both suspension types (Figure 5.14). The achromatic and yellow data are closely aligned suggesting the yellow pigments had little impact over their common range (up to around 250 NTU). Although non-linear overall, the lower ranges of the data appear to be linear. Applying linear regression to a truncated range of the combined dataset (achromatic and yellow data, capped at approximately 50 NTU) results in an $R^2$ value of 0.9987 (Figure 5.15).

### 5.4.2 Microsphere concentration and aggregate luminance

The relationship between aggregate luminance and microsphere concentrations was found to follow a broadly similar response to that of instrumental turbidity measurements. Since the camera parameters and nominal illuminants are constant across the entire dataset, a comparative assessment of the surrounds and suspension types can be made (Figure 5.16). The booth surround is seen to have a substantial effect on the luminance of the samples with the light surround values approximately double those of the dark surround. This is likely caused by the higher reflectivity of the light walls causing more light to enter the sample at all angles enclosed by the walls. At higher levels the compression effects in the yellow suspension appear larger than that for the achromatic suspension (higher range data only available for the light surround). This is likely due to cumulative scatter effects associated with the yellow pigments. In a similar fashion to the treatment of the turbidity data, linear regression was applied to a truncated range of the combined achromatic and yellow data for each of the booth surround types (Figure 5.17). Again the data was strongly linear over the shortened range (dark surround, $R^2$=0.9977; light surround, $R^2$=0.9823).

### 5.4.3 Relationship between aggregate luminance and instrumental turbidity

The results outlined above would suggest a strong linear relationship between aggregate luminance and turbidity in the lower ranges; this was confirmed by linear regression (Figure 5.18 and Table 5.11 for details). The increased illumination of the light surround results in a larger intercept value for suspensions in the light surround compared to the dark surround. In contrast, the light surround appears to depress the slope compared to the dark surround. The dark surround results generated the highest $R^2$ value; this would be expected since the dominant overhead illumination in the dark surround better emulates the optics of the nephelometer.

**Table 5.11 Linear regression results for aggregate luminance versus instrumental turbidity for combined achromatic and yellow suspension data. Range truncated to around 50 NTU.**

| Surround | Slope | Intercept | $R^2$ |
|---|---|---|---|
| Dark | 0.28 | 5.2 | 0.9978 |
| Light | 0.21 | 29.4 | 0.9813 |

### 5.4.4 Channel sensitivity to yellow pigments

The aggregate luminance values previously discussed are weighted sums of the output from individual RGB channels and these channel values are heavily influenced by the spectral response of the corresponding filter in the camera. In an effort to determine which channel might offer the optimum linearity with respect to a mixed sample set (i.e. samples containing achromatic and yellow microspheres), the individual RGB channel responses were subjected to linear regression. Figure 5.19

shows the results for the dark surround where the yellow and achromatic suspension data has been combined and shown over the previously used truncated range (to around 50 NTU). It can be seen that the blue channel response suffers from substantial oscillations due to its sensitivity to the yellow pigments and the mix of yellow and achromatic data points. The green and red channels are relatively less affected with the red channel offering the most robust response to mixed yellow and achromatic samples.

### 5.4.5 Summary and conclusions

The relationship between turbidity and microsphere concentration was found to be a non-linear, compressive function over the range tested. A different positive exponential response was reported by Carrasco and Siebert using microsphere diameters of 0.769 µm over a similar turbidity range[3]. However, an examination of a more complete set of data from the underlying study shows that both compressive and expansive responses were found depending on the configuration of the nephelometer[54]. It is likely that a combination of different particle diameters along with inconsistencies in turbidity measurements associated with differing nephelometer designs has contributed to the difference in the results. A non-linear response was also found for the relationship between image based luminance data and microsphere concentration.

When the data range was truncated to around 50 NTU, both turbidity and aggregate luminance showed a strong linear relationship with microsphere concentration. Although the relationship between instrumental turbidity and aggregate luminance was linear below 50 NTU, the slopes and intercepts varied between surround types. The linearity of the aggregate luminance to turbidity values for combined achromatic and yellow datasets was improved by removing the blue channel data from aggregate luminance. These are novel results and no other relevant published data is available for comparison.

The linearity of the turbidity/luminance response suggests that digital photography in combination with appropriate overhead lighting may offer a means of emulating the function of traditional nephelometric instruments. Practical use of this method would need a means of establishing a calibration between the device dependent RGB colour channel values that underpin the luminance measures and the nephelometric scale based on formazin.

The potential of image based turbidity measurement is investigated further in Chapter 7.

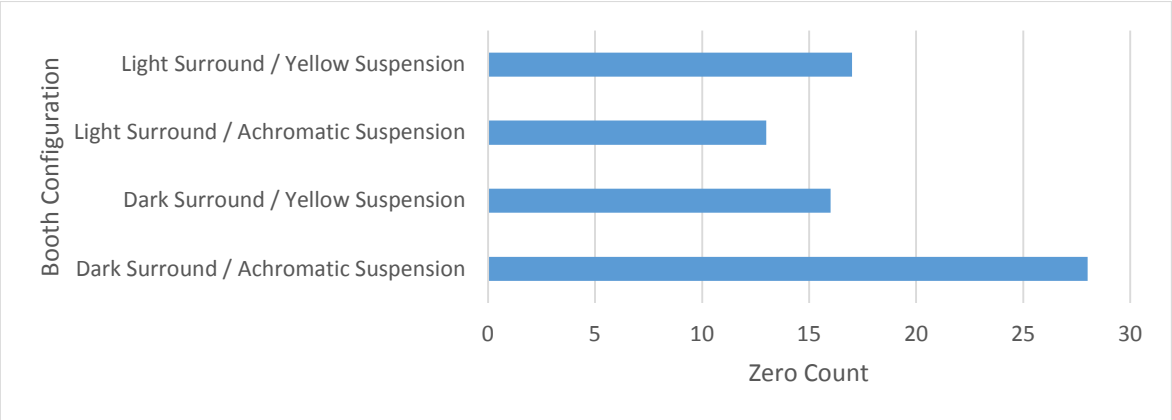


**Figure 5.1 Distribution of zero evaluations across each of the experimental configurations.**

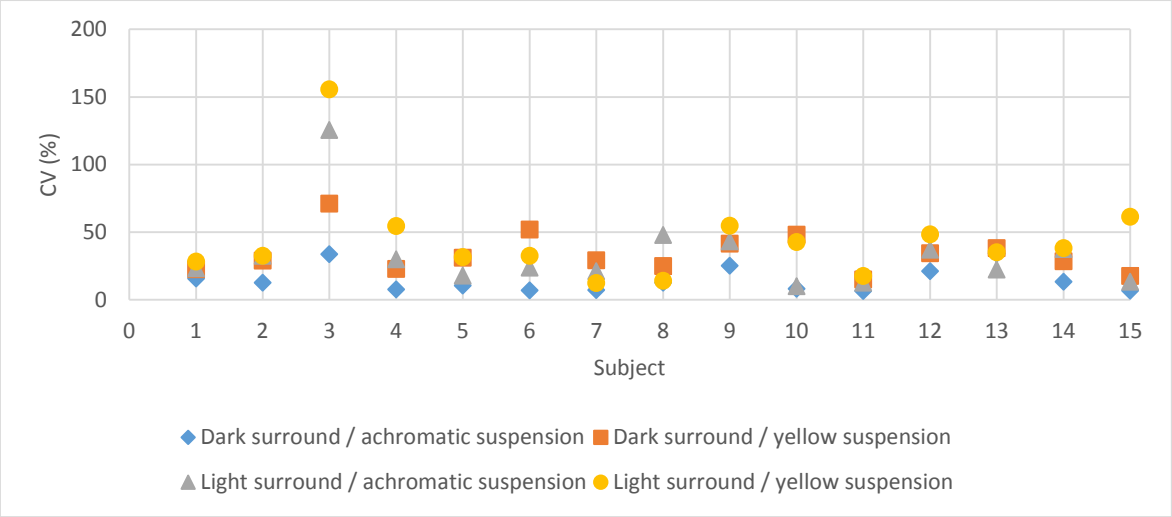


**Figure 5.2 Coefficient of variance (CV) values by subject by experimental configuration.**

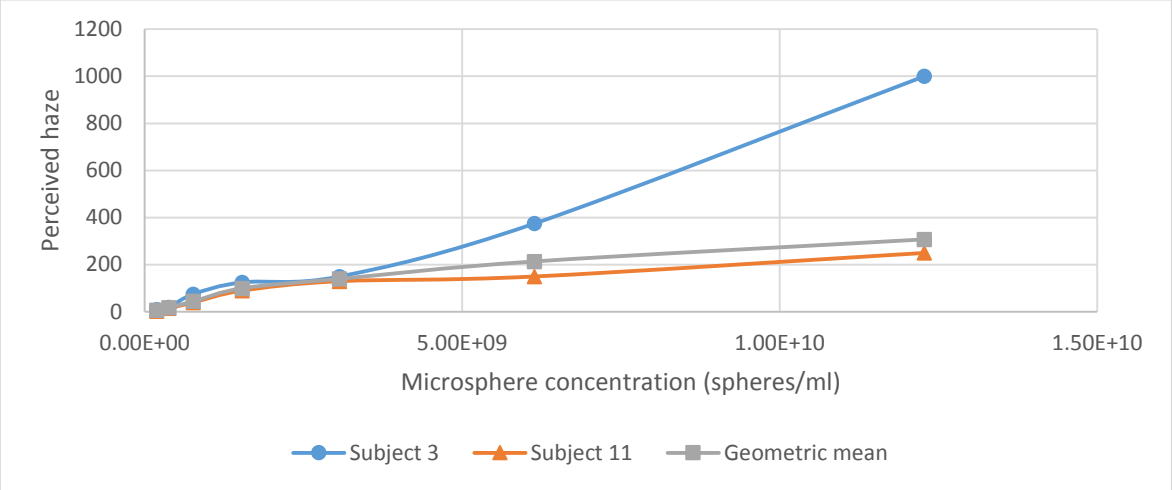


**Figure 5.3 Subject 3 and Subject 11 evaluation plotted with overall geometric mean (light booth / yellow suspension). Cubic spline fit to illustrate trend.**

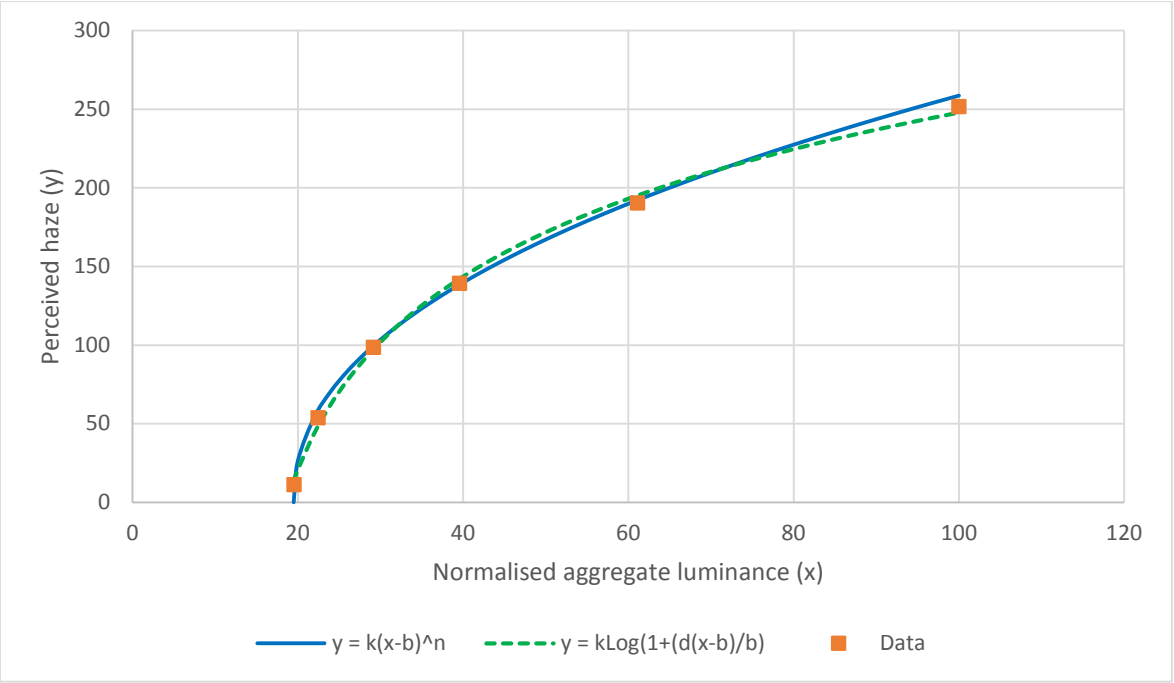


**Figure 5.4** **Perceived haze versus aggregate luminance for dark surround and achromatic suspensions. Best fit log and power functions shown; power function provides optimum fit with an $R^2$ value of 0.9991.**

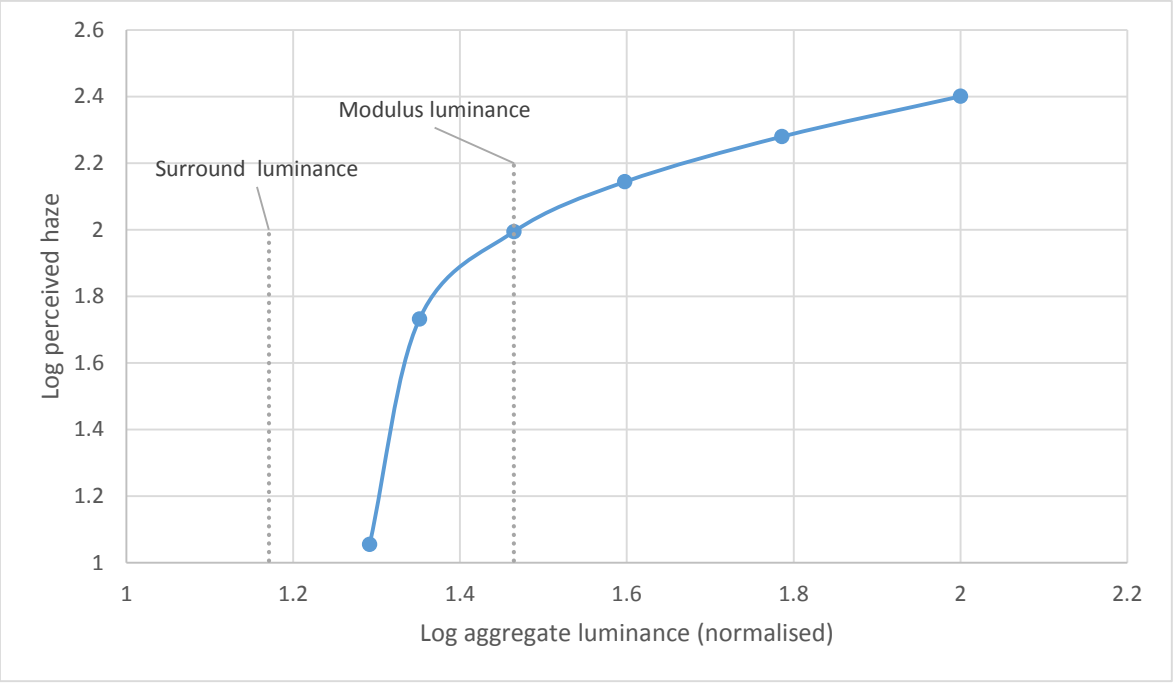


**Figure 5.5** **Log perceived haze versus log aggregate luminance for dark surround and achromatic suspensions. Cubic spline fit to illustrate trend.**

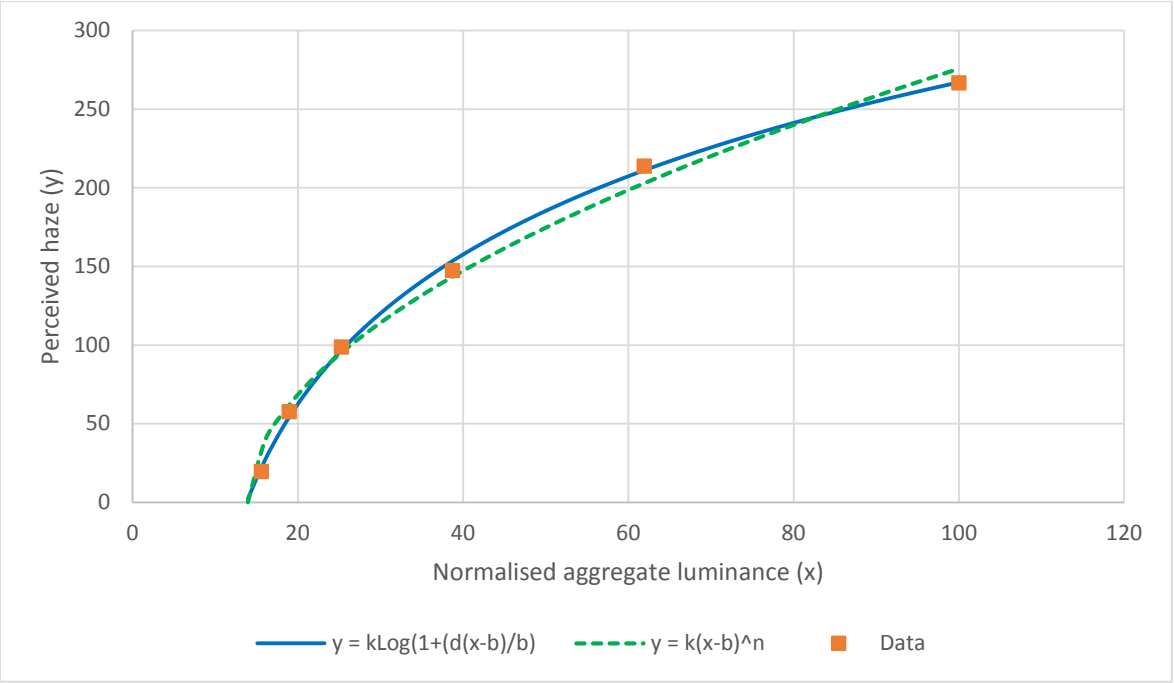


**Figure 5.6** **Perceived haze versus aggregate luminance for dark surround and yellow suspensions. Best fit log and power functions shown; log function provides optimum fit with an $R^2$ value of 0.9988.**

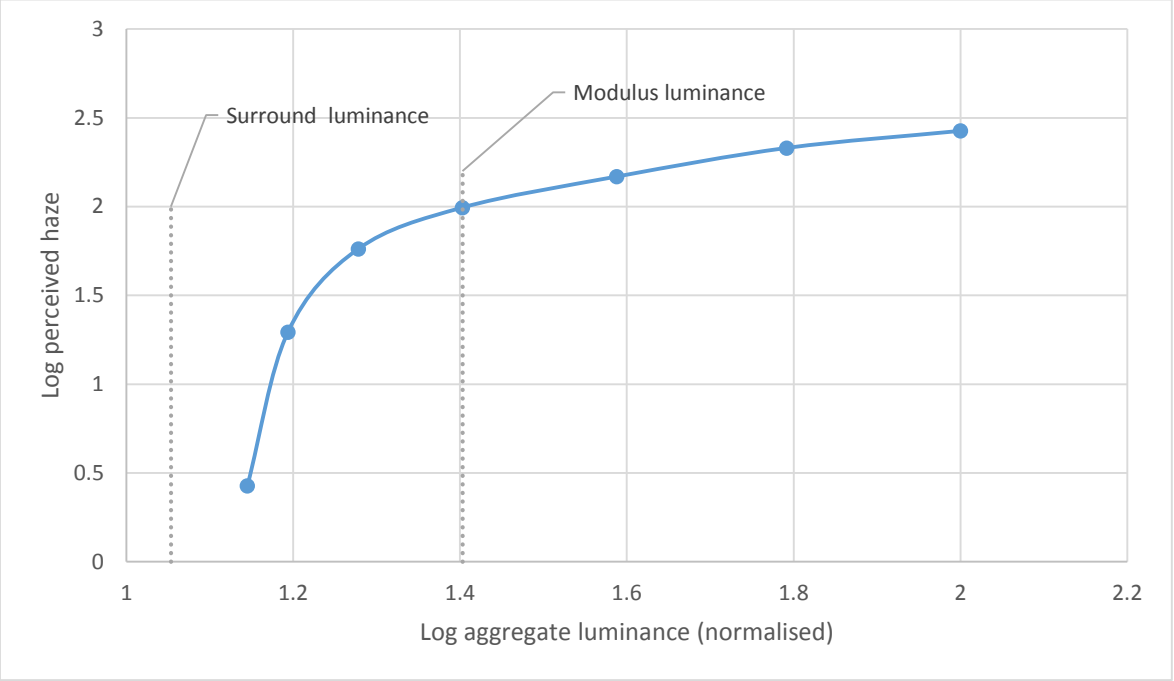


**Figure 5.7** **Log perceived haze versus log aggregate luminance for dark surround and yellow suspensions. Cubic spline fit to illustrate trend.**

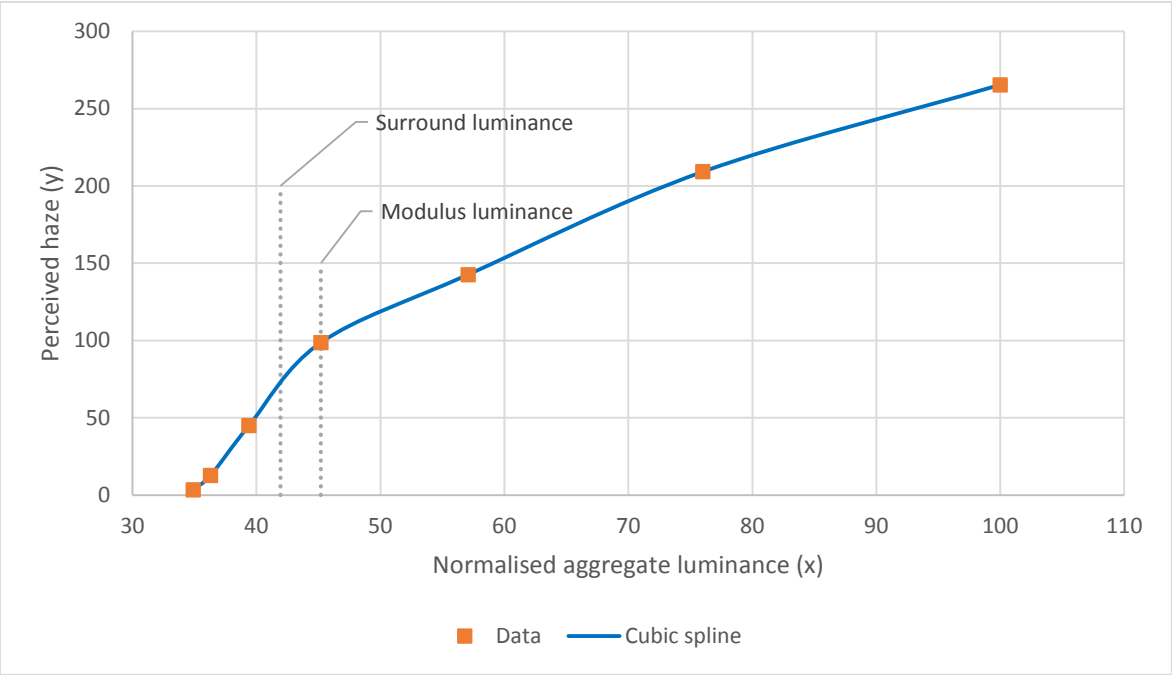


**Figure 5.8** **Perceived haze versus aggregate luminance for light surround and achromatic suspensions. Cubic spline fit to illustrate trend.**

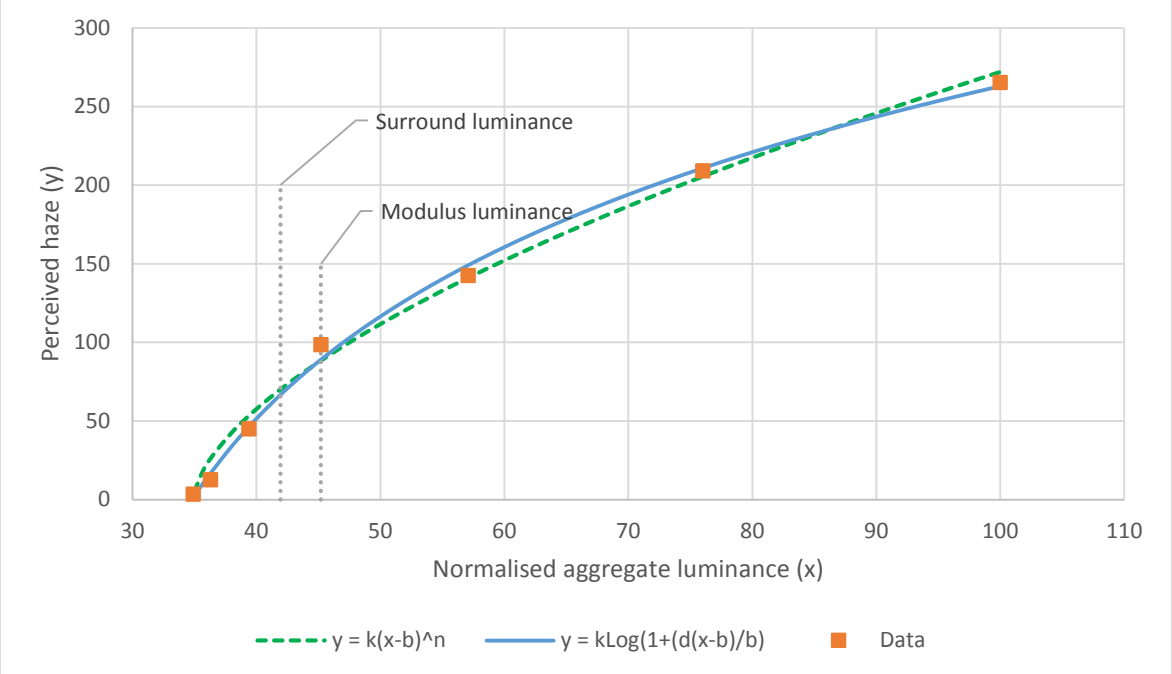


**Figure 5.9** **Perceived haze versus aggregate luminance for light surround and achromatic suspensions. Best fit log and power functions shown; log function has highest $R^2$ value of 0.9972 but neither function fits the structure of the data.**

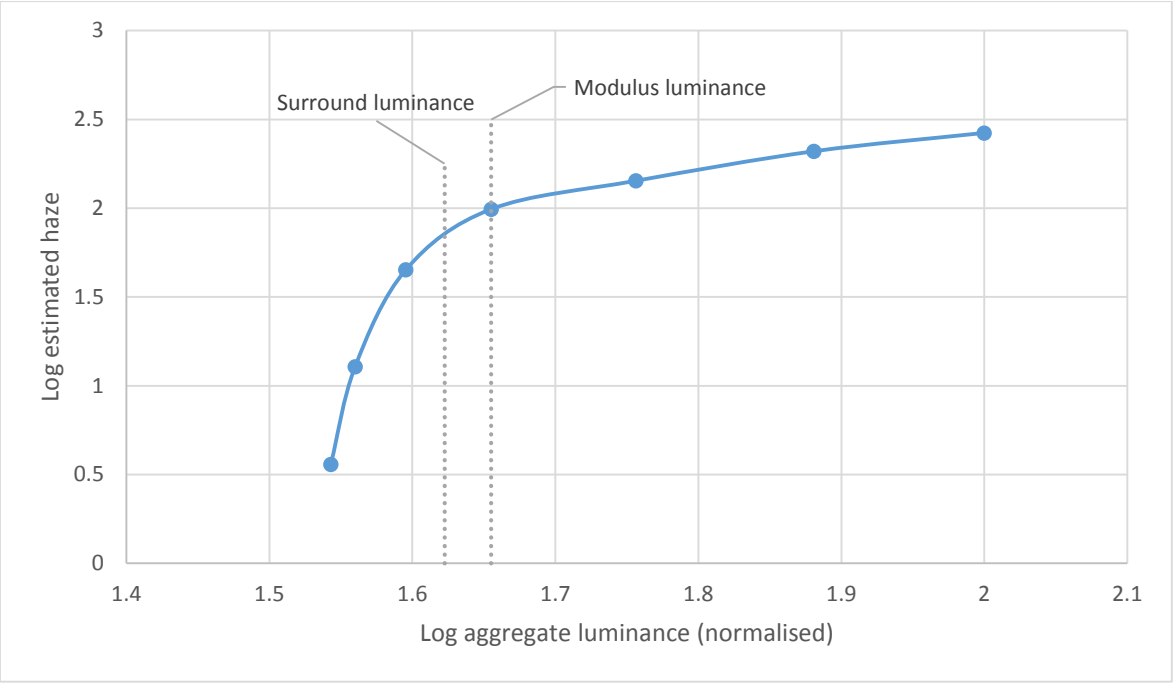


**Figure 5.10 Log perceived haze versus log aggregate luminance for light surround and achromatic suspensions.**

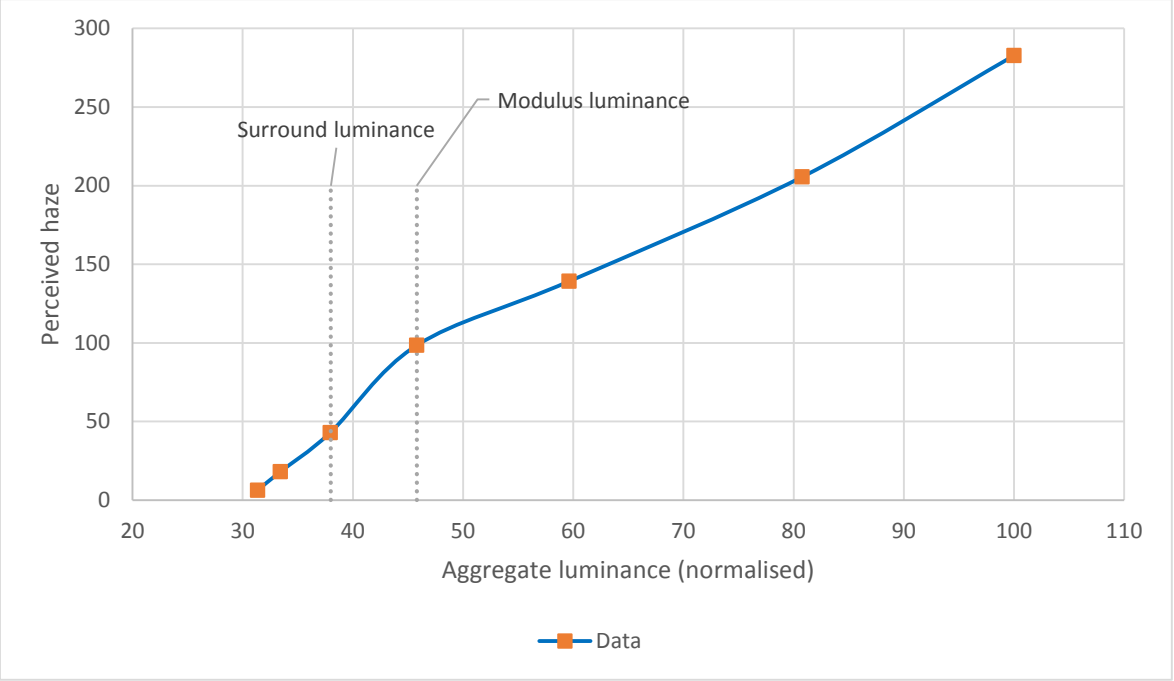


**Figure 5.11 Perceived haze versus aggregate luminance for light surround and yellow suspensions. Cubic spline fit to illustrate trend.**

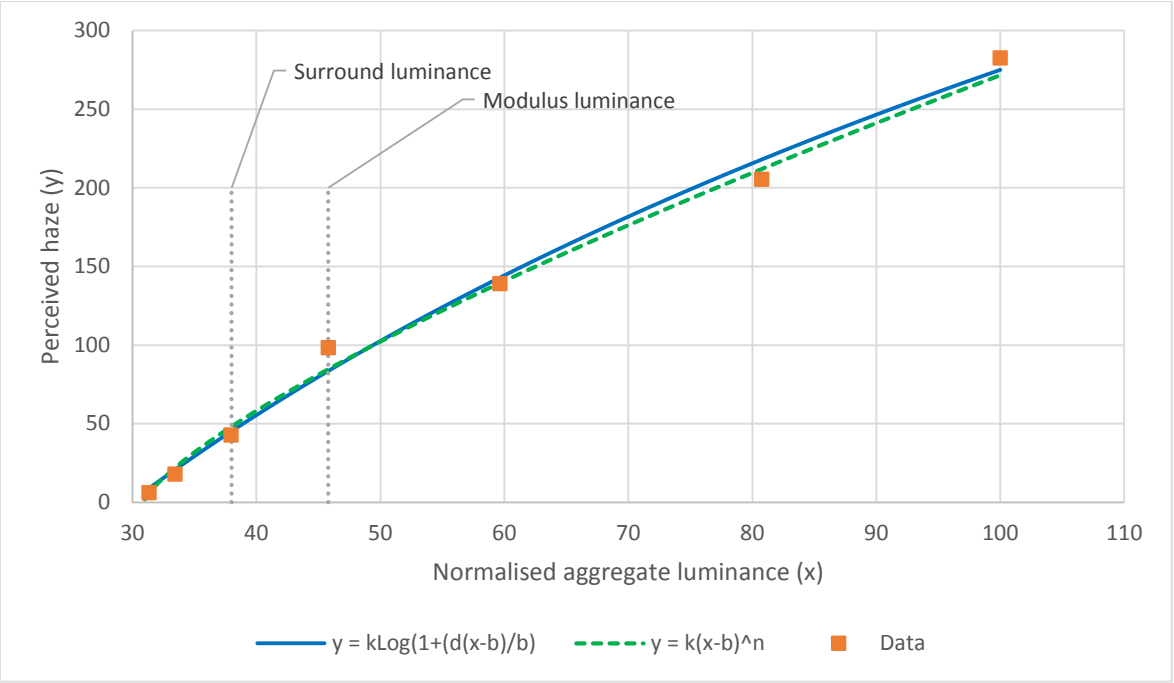


**Figure 5.12 Perceived haze versus aggregate luminance for light surround and yellow suspensions. Best fit log and power functions shown; log function has highest $R^2$ value of 0.9944 but neither functions fit the structure of the data.**

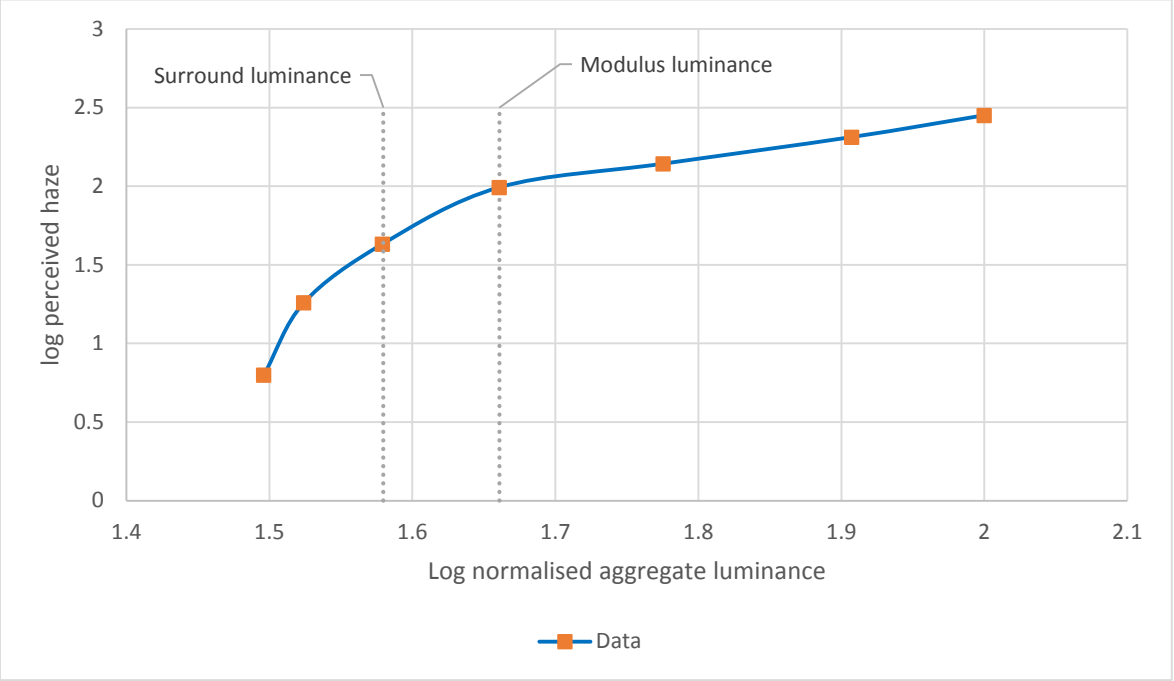


**Figure 5.13 Log perceived haze versus log aggregate luminance for light surround and yellow suspensions.**

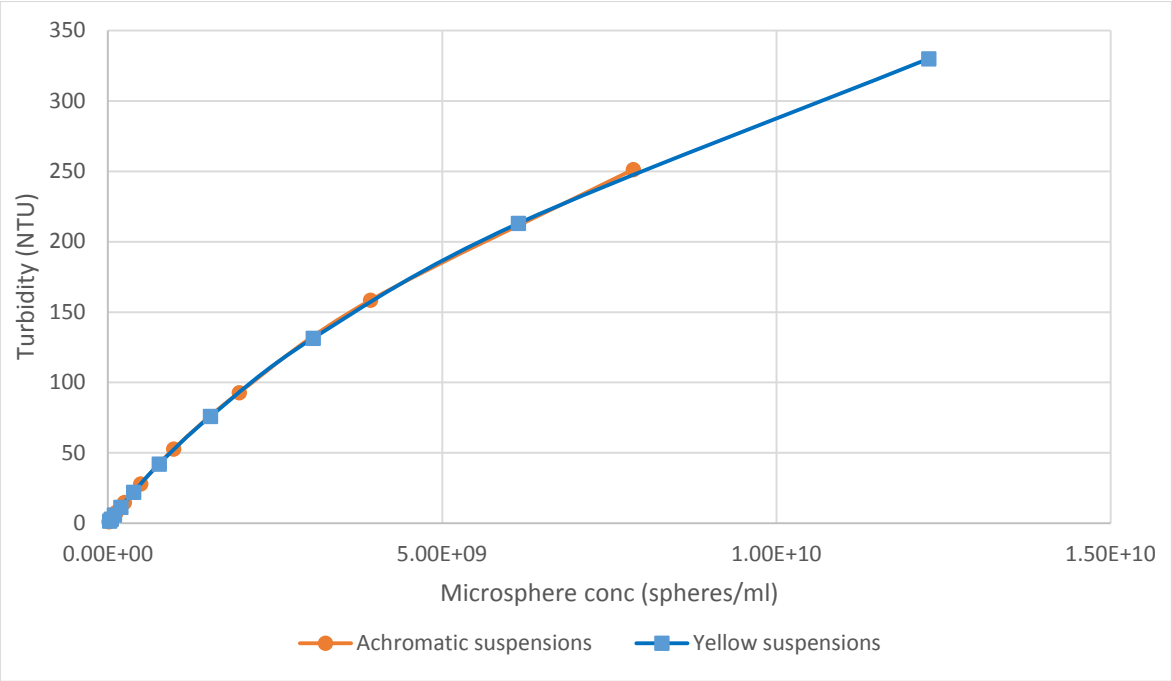


**Figure 5.14 Instrumental turbidity versus microsphere concentrations for achromatic and yellow suspensions (cubic spline fit for trend illustration).**

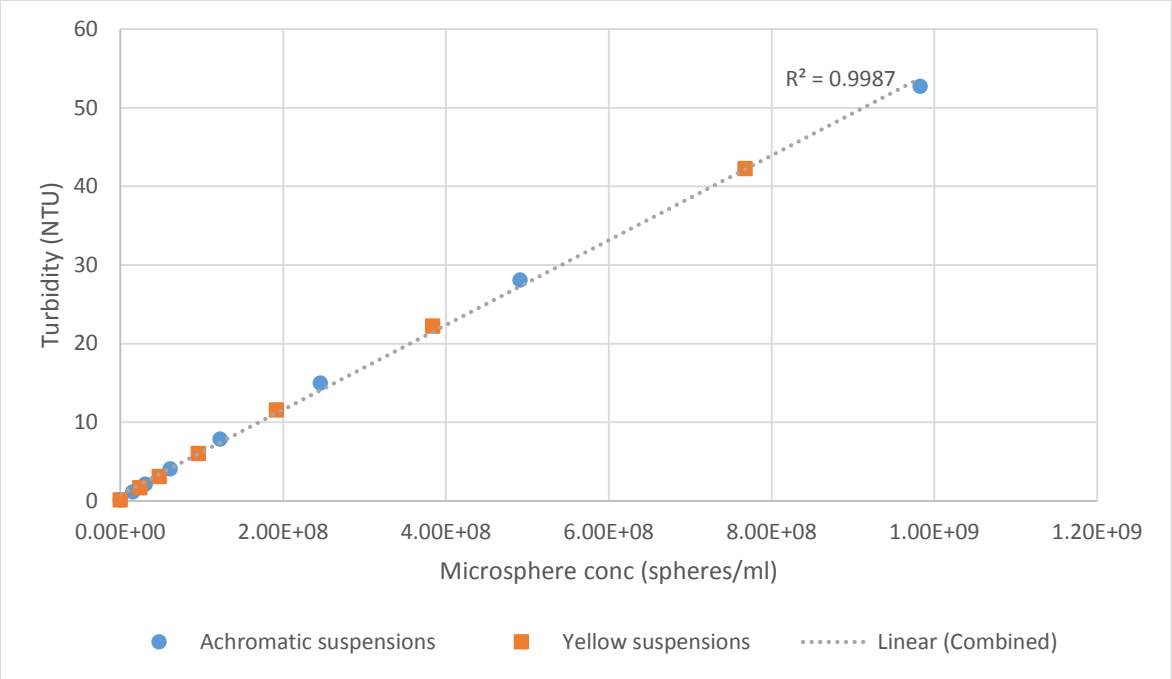


**Figure 5.15 Instrumental turbidity versus microsphere concentrations for achromatic and yellow suspensions. Range truncated to the equivalent of approximately 50 NTU. Linear regression applied to combined achromatic and yellow dataset.**

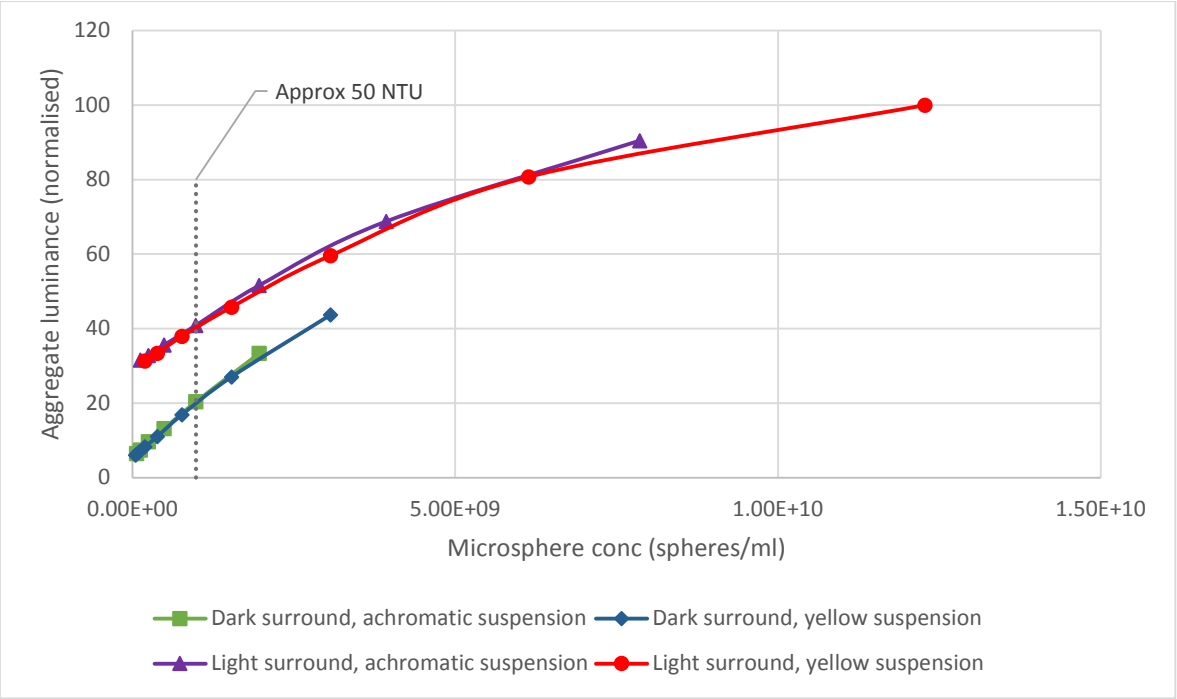


**Figure 5.16** **Aggregate luminance versus microsphere concentration for all surrounds and suspension types (luminance values normalised to highest luminance sample in light surround). Cubic spline fit for trend illustration.**

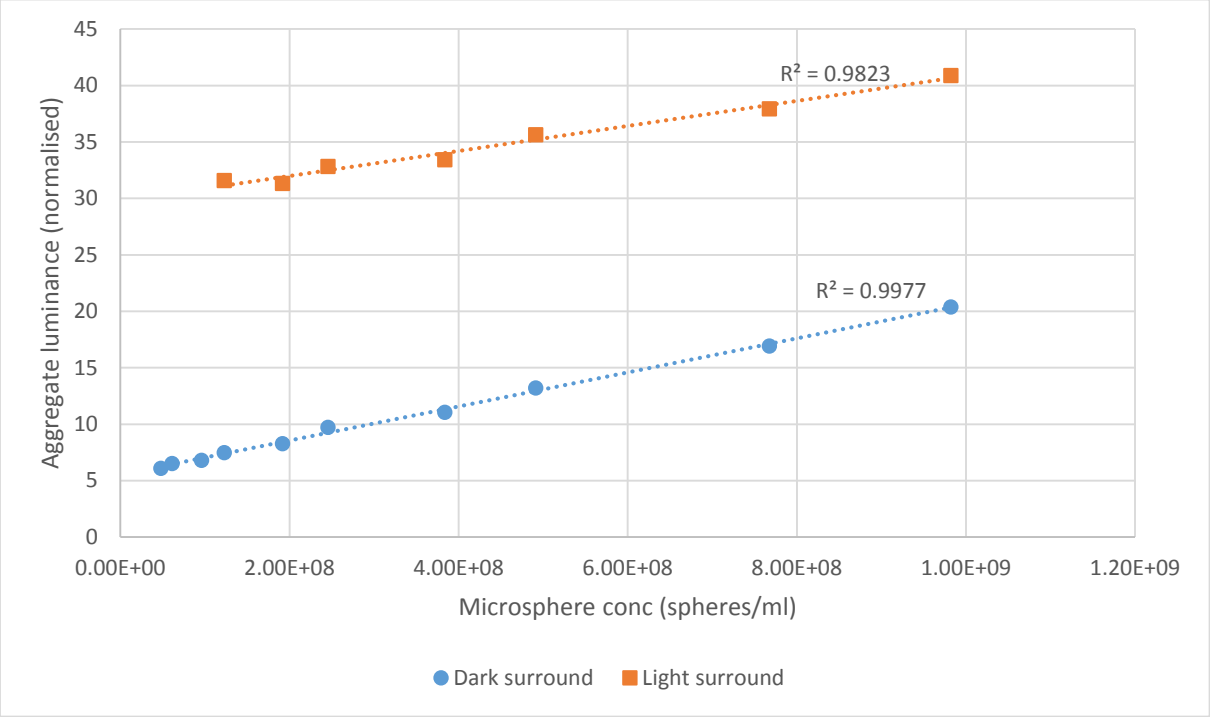


**Figure 5.17** **Aggregate luminance versus microsphere concentrations for dark and light surrounds. Range truncated to the equivalent of approximately 50 NTU. Linear regression applied to combined achromatic and yellow datasets for each surround type.**

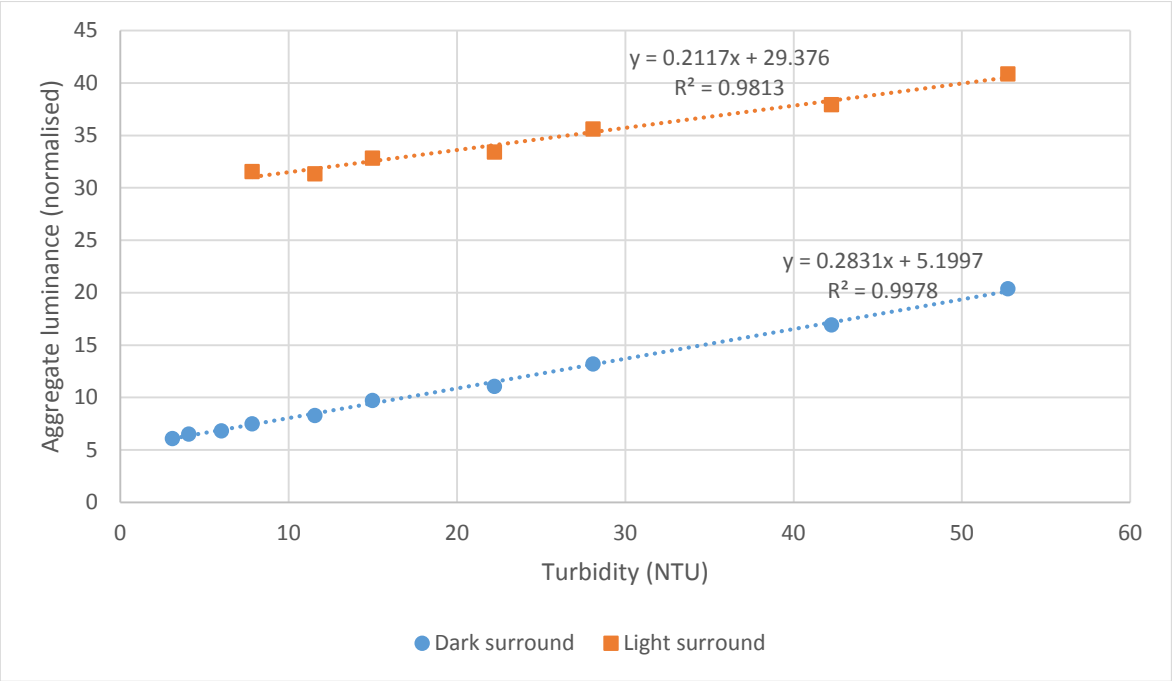


**Figure 5.18** **Aggregate luminance versus instrumental turbidity for dark and light surrounds. Range truncated to the equivalent of approximately 50 NTU. Linear regression applied to combined achromatic and yellow datasets for each surround type.**

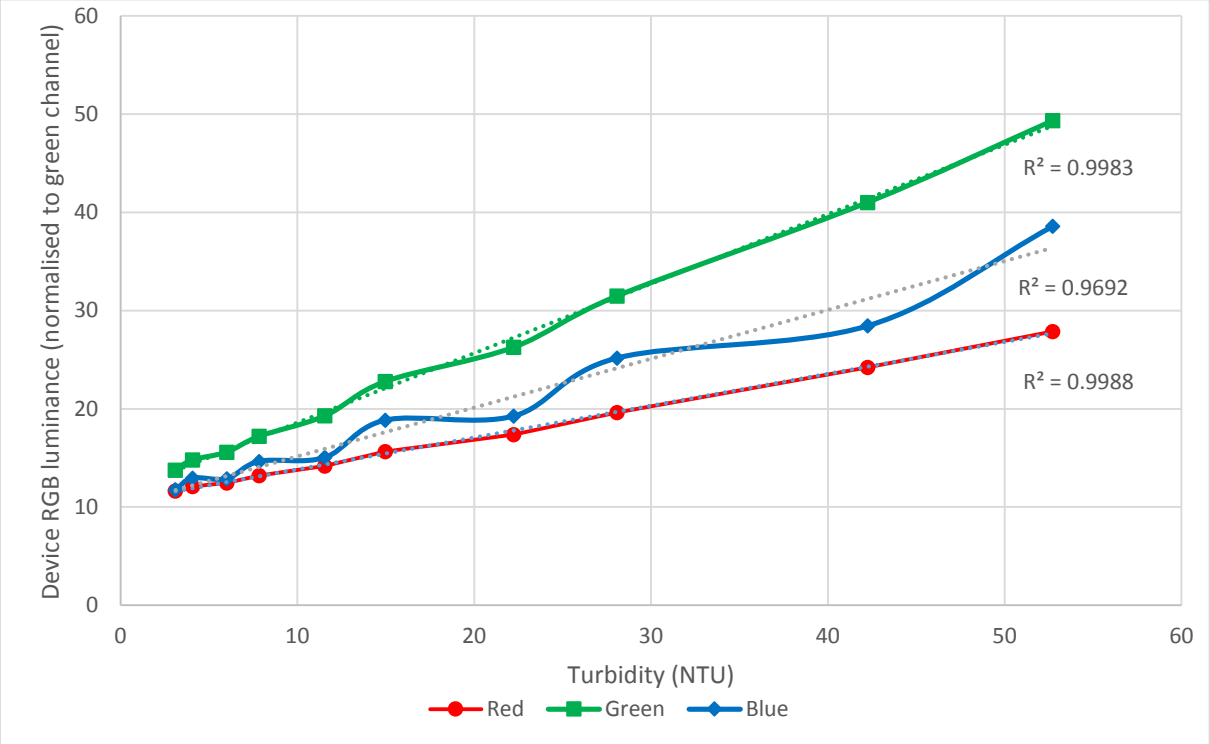


**Figure 5.19** **Linear regression results for individual RGB channel luminance versus turbidity in dark surround. Achromatic and yellow suspension data combined to highlight channel sensitivity. Cubic spline fit added to data to illustrate trend.**

# Chapter 6

# Relationships between Haze Intensity Perception and Nephelometric Measurements in Low Haze Chardonnay Wine Samples

## 6.1 Overview and objectives

The relationship between human haze perception and instrumental measures in yellow polymer microsphere solutions had been explored over a wide range of turbidities; 1.6 Nephelometric Turbidity Units (NTU) to 330 NTU (see Chapter 5). Haze intensity perception was found to be best described by a logarithmic function of the general form of Equation 3.7. Although the experimental data was well described by the log function, the data was truncated at 3.1 NTU due to zero evaluations by the subjects for the lowest concentration levels. The overall goal of this study was to extend the previous work to commercial wine samples at a lower turbidity range and an increased data resolution. Turbidity was to be used as the independent variable as opposed to image luminance. Subject evaluations were to be conducted under conditions closely resembling those encountered when wine is being consumed in the presence of a dark surround. Test materials were to consist of haze induced commercial grade Chardonnay wine samples. The specific objective was to determine if any of the generalised models defined in Section 3.5.2 would fit the human response to low turbidity wine samples viewed in a dark surround.

## 6.2 Materials and methods

The information given below should be read in conjunction with the general methodology Section found in Chapter 3.

### 6.2.1 Sample preparation, ranges and stability monitoring

The results of a related threshold detection experiment (see Section 4.3.3) had shown that the group threshold for Chardonnay wine samples in a dark surround was around 1.2 NTU. The sample preparation plan was designed to create a range of 10 samples from just above group threshold up to around 25 NTU in concentration steps of 1.4. The modulus was to be created using a 50%/50% mix of samples 5 and 6 to provide an intermediate reference.

Wine stock exhibiting a high degree of visual haze was prepared by immersing 1 L volumes of Chardonnay wine in a temperature controlled heat bath at 90˚C for a period of two hours; this

formed a concentrate base for dilutions. The haze induced wine samples were prepared by diluting the high haze wine stock with untreated wine (Table 6.1).

Once the samples had been prepared, pilot work showed that the lowest three sample levels in were not easily discriminated and therefore could not be ranked and scaled under the intended viewing conditions. This supported previous work that suggested haze perception threshold levels may not be a good guide to the lowest levels at which haze could be scaled (see Section 5.3.1). It was decided to exclude the lowest two levels from the study resulting in an experimental sample range of 2.3 NTU to 25.5 NTU with a modulus of 4.2 NTU.

Since the nominal sample range extended above 10 NTU, the possibility of excessive sample instability during subject evaluations was anticipated (see Section 3.4.3). To ensure that errors in the average turbidity values were minimised, a regime of regular turbidity measurements was employed over the course of the experiment (each sample was measured on 4 occasions over the course of the experimental period; an elapsed time of 9 hours). The resulting average turbidities and associated Standard Deviation values are shown in Table 6.1. It can be seen that relatively high instabilities are observed in samples above 9.3 NTU.

**Table 6.1 Details of Chardonnay haze induced sample suspensions. Average turbidity values and standard deviations derived from four individual measurements taken during the experimental day (9 hours).**

| Chardonnay samples | Average turbidity (NTU) | Standard deviation |
|---|---|---|
| Untreated/Blank[1] | 0.4 | NA |
| 1[1] | 1.4 | NA |
| 2[1] | 1.7 | NA |
| 3 | 2.3 | 0.03 |
| 4 | 2.9 | 0.02 |
| 5 | 3.6 | 0.08 |
| Modulus | 4.2 | 0.09 |
| 6 | 4.9 | 0.11 |
| 7 | 6.4 | 0.05 |
| 8 | 9.3 | 0.04 |
| 9 | 14.7 | 0.30 |
| 10 | 25.5 | 0.55 |

[1] Not used in sensory evaluations.

### 6.2.2 Sensory booth configuration

Sensory booths were equipped and illuminated as previously described. A single booth with a dark surround was used.

### 6.2.3 Sample handling and presentation

A sample set consisting of eight ISO glasses (each containing a 90 m sample) were prepared along with a mid-range reference modulus containing the same volume. Each glass was labelled with a random three digit code.

### 6.2.4 Sensory evaluations of haze intensity

#### Overview

The previous haze scaling study had employed two-fold dilution steps and this resulted in generally acceptable levels of discrimination errors in the sensory data (see Section 5.3.1). However, the yellow suspensions/dark surround combination had led to the highest error rate of 4.8%. Since this configuration corresponded most to this study and the concentration steps were much compressed (a factor of 1.4 versus 2.0), increased error rates were anticipated if the same scaling method was used (i.e. Magnitude Estimation).

In an attempt to improve discrimination error rates in scaling studies Kim and O'Mahony[118] proposed a modification to the traditional Magnitude Estimation technique whereby subjects are required to sequentially place samples on a physical scale so that the samples are simultaneously ranked and scaled. A key aspect of the technique (known as Rank-Rating) is that subjects can re-evaluate and re-position any test sample as they progress through the sample set. The intention is to minimise errors due to subjects forgetting the values assigned to previously rated samples. Although other studies that have used Rank-Rating appear limited to taste experiments[118,124], it was thought that a suitably modified version of Rank-Rating might be used in this study as a means of limiting errors caused by the compressed sample range. Physical constraints in the sensory booths meant that it was not possible to employ a physical scale exactly in the manner of Kim and O'Mahony[118]. The procedure was modified so that subjects conducted a ranking task that then informed a second stage Magnitude Estimation scaling task. The "scale as a group" principle was retained whereby subjects were allowed to rescale/reposition samples freely.

The eight individual evaluation sessions were scheduled for a duration of one hour and were conducted over an elapsed nine hour period.

#### Subjects

Subjects for this study consisted of eight volunteers (staff, faculty members or post graduate students) from Lincoln University, New Zealand. All subjects self-reported normal or corrected vision. The subject pool included three males and five females all over the age of 18. All of the subjects had been involved in a previous study that used Magnitude Estimation to evaluate haze in polymer

microsphere suspensions. The subjects were not given any monetary or other incentive to participate in the study.

**Evaluation procedure**

On arrival the subject was directed to the sensory booth area and given a briefing that was read from a pre-written script (Appendix C.4). It was explained that a modified version of the previously employed Magnitude Estimation technique was to be used to evaluate the intensity of haze in wine samples. In particular, the process of evaluating the samples as a group was explained in some detail. This pre-evaluation briefing took approximately five minutes; a reasonable time to allow the subject's vision to adapt to the sensory room.

In advance of the subject arriving, all eight samples had been arranged in a "clock like" circle in the sensory booth. The order of the samples in the circle was changed for each subject in accordance with an 8x8 Latin Square structure so that each of the eight samples appeared in each initial position and in every other position an equal number of times before re-using the presentation order.

Once seated in the booth, the subject was asked to rank the samples in a row (order was alternated by subject: ascending haze or descending haze). On completion of the ranking task, the subject was then shown the modulus and told it had a haze value of 100 against which all test samples were to be judged (no zero evaluations were allowed, no upper limit to the evaluations). The subject was then asked to work from left to right verbally assigning an intensity value to sample. The hand with which the modulus was held was alternated by subject and the scaling evaluation was performed centrally in the booth at eye-level. The experimenter recorded the evaluations on individual scoring sheets. Once the subject had rated the intensity of all samples, the scores were placed in front of the samples and the subject was given the opportunity to freely review and re-evaluate samples as they wished.

The experimenter remained present at the sensory booth throughout the evaluation process. No time limits were set for any aspect of the evaluation.

### 6.2.5 Data analysis and curve fitting

The intensity of haze perception was quantified by taking the geometric mean of the final magnitude estimates given by the subjects resulting in a single geometric mean value for each of the samples.

Non-linear regression techniques were used to fit the experimental data to various general brightness/lightness models as previously described (Section 3.5.2). Instrumental turbidity (NTU) was used as the independent variable and the geometric means of subjects' evaluations as the dependent variable.

## 6.3 Results and discussion

### 6.3.1 Review of subjects' evaluation data

#### Discrimination errors

In Magnitude Estimation studies the steps between stimuli are generally chosen such that each step is easily discernable by the subjects[105]. However, this study purposely compressed the inter-stimuli steps in order to capture a relatively short, high resolution range and this raised the risk of increased discrimination errors. A discrimination error is observed when a subject gives a rating to higher concentration that is less than or equal to a rating assigned for a lower concentration.

A substantially increased error rate of 18.8% was observed in comparison to a rate of 4.8% for yellow polymer microsphere suspensions in the same surround (see Section 5.3.1). Careful examination of the data found clustering of errors either side of the modulus with 75.0% of the total errors located in this area. All but one of the subjects made discrimination errors around the modulus. It is worth noting that fixing the modulus mid-way between two samples compressed the range around the modulus in excess of the design dilution factor of 1.4. It is likely that the subjects could not easily discern Samples 5 and 6 from the modulus due to the small step size at these levels.

#### Inter-subject variation

Inter-subject variation was calculated using the Coefficient of Variance (CV) method previously described using the geometric mean as the base reference; the overall arithmetic mean of the subjects' CV values was found to be 23.0%. This is almost equivalent to the CV value found for yellow polymer microsphere suspensions in the same surround (22.4%). However, as noted previously, the most comparable previous study was performed by Martin et al. and a mean value CV of 17% was found for lightness[34]; the CV values found for this study are therefore somewhat high.

Individual CV values were examined for bias or unusual ratings; no unusual outliers were observed (see Figure 6.1).

#### Ranking and rating inconsistencies

Three of the eight subjects encountered conflicts between their chosen ranking order and the intensity score they gave to the samples. The subjects re-examined the samples concerned and, in each case, chose to retain the score assigned. The conflicts were likely caused by the asymmetric nature of the lighting resulting in a bias in the ranked samples when arranged in a row. Since the subjects gave priority to their scores rather than the nominal rank, the resulting impact on the data is not thought to be significant.

### 6.3.2 Data fitting and optimum model

The relationship between perceived haze and turbidity values was found to exhibit something of a bi-segment response. However, the discrimination errors associated with the sample levels adjoining the modulus make any determination of the response uncertain. Non-linear regression results are summarised in Table 6.2; data and best fit curves for the entire dataset are shown in Figure 6.2. Although the $R^2$ value for log function of Equation 3.7 is superior, neither fit is perfect.

Whilst recognising the uncertainty in the region of the modulus, it does appear that a two segment response is visible in the log-log plot with the modulus acting as the boundary (Figure 6.3). The lower segment is likely linear and appears to conform to a power function with an exponent of approximately 1.4. The upper segment is not linear and does not appear to correspond to a power function.

**Table 6.2 Results for non-linear regression of perceived haze versus turbidity (NTU) for haze induced Chardonnay wine samples. Log and power models with highest $R^2$ values shown.**

| Model | Form | k | b | d | n | $R^2$ |
|---|---|---|---|---|---|---|
| Eqn 3.7 | $y = k\log(1 + d(x-b)/b)$ | 1092.8 | 0.31 | 1.78E-02 | NA | 0.9961 |
| Eqn 3.3 | $y = k(x-b)^n$ | 44.3 | 1.11 | NA | 0.71 | 0.9949 |

### 6.3.3 Summary and conclusions

The intention of this study was to explore a lower range at a higher resolution than that which had been examined in the study of yellow microsphere suspensions. Although data was successfully collected, the small step size around the modulus and associated uncertainty in the data undermined the results somewhat. The clustering of discrimination errors around the modulus level was unfortunate since the microsphere studies had suggested that this is the region in which discontinuities caused by induction effects were most likely to be observed when the modulus is present during evaluation. The discrimination errors around the value of the modulus served to obscure this area of interest. Additionally, the inability of subjects to discern differences at the lower end of the proposed range forced the removal of the two lowest samples from the study. This further undermined the aim of studying the lowest part of the range.

The apparent inability of the subjects to discern differences between samples lower in the turbidity range despite strong evidence from the threshold studies that they could detect haze at those levels, presents a serious challenge in studying the haze intensity response at low turbidity values. A similar problem was found in the equivalent microsphere scaling study where the subjects had a tendency to assign zero values at low supra-threshold levels (see Section 5.3.1). This is unfortunate since small

value turbidity differentials close to the threshold are of most interest to the winemaker in protein stabilisation tests where stability differentials of 2 NTU are commonly cited[1].

The results of this study indicate that the response of haze induced Chardonnay wine samples are broadly similar to that observed for the yellow microsphere suspensions and the detailed conclusions in Section 5.3.6 are generally supported. Should it be necessary to predict the human response to a Chardonnay wine sample exhibiting haze based on turbidity measurements, a power function would appear most appropriate using 0.7 for the exponent value (assuming viewing conditions and turbidity range similar to those used in this study).

Any future work should ensure that the physical value/concentration of the modulus does not conflict with the overall concentration scheme for the test samples; doing so risks unacceptably high discrimination errors at levels adjoining the modulus. As previously noted, the modulus appears to have a strong inductive effect and measures should be taken to ensure that it is not present whilst evaluations are being made (unless the effect of paired comparisons are to be studied).

It is not possible to comment fully on the ability of the modified Rank-Rating technique to reduce discrimination errors due to the difficulties caused by the small step size around the modulus. However, if the errors around the modulus are excluded, the error rate drops to 2.7% and this is low in comparison to the equivalent rate of 4.8% found in the microsphere study despite the severely compressed range. An additional area of concern is the use of a row layout since this can lead to scoring inconsistencies if the illumination levels along the row are not constant (e.g. samples may appear brighter simply due to their position in the row if the lighting is not adequately diffuse). Any future visual study that employs a Rank-Rating technique based on a row layout should ensure that the area is uniformly illuminated.

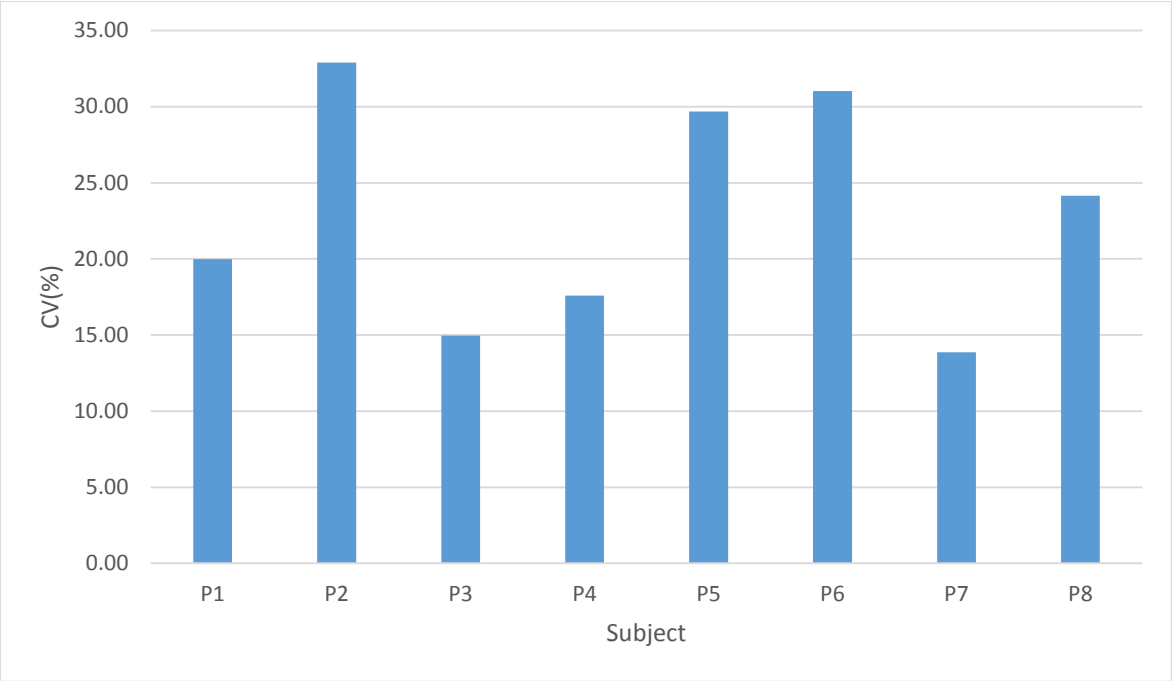


**Figure 6.1** **Coefficient of variance (CV) values by subject for haze induced Chardonnay samples using modified Rank-Rating technique.**

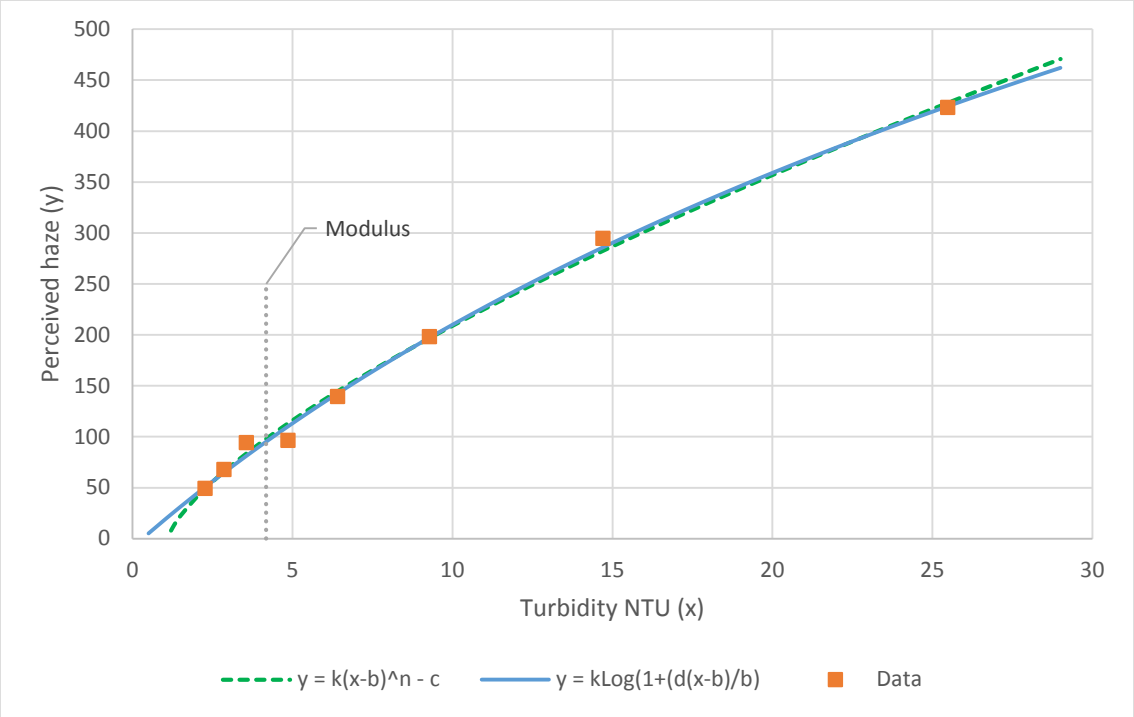


**Figure 6.2** **Perceived haze versus instrumental turbidity for haze induced Chardonnay wine samples. Best fit log and power functions shown; power function offers optimum structural fit with an $R^2$ value of 0.9949.**

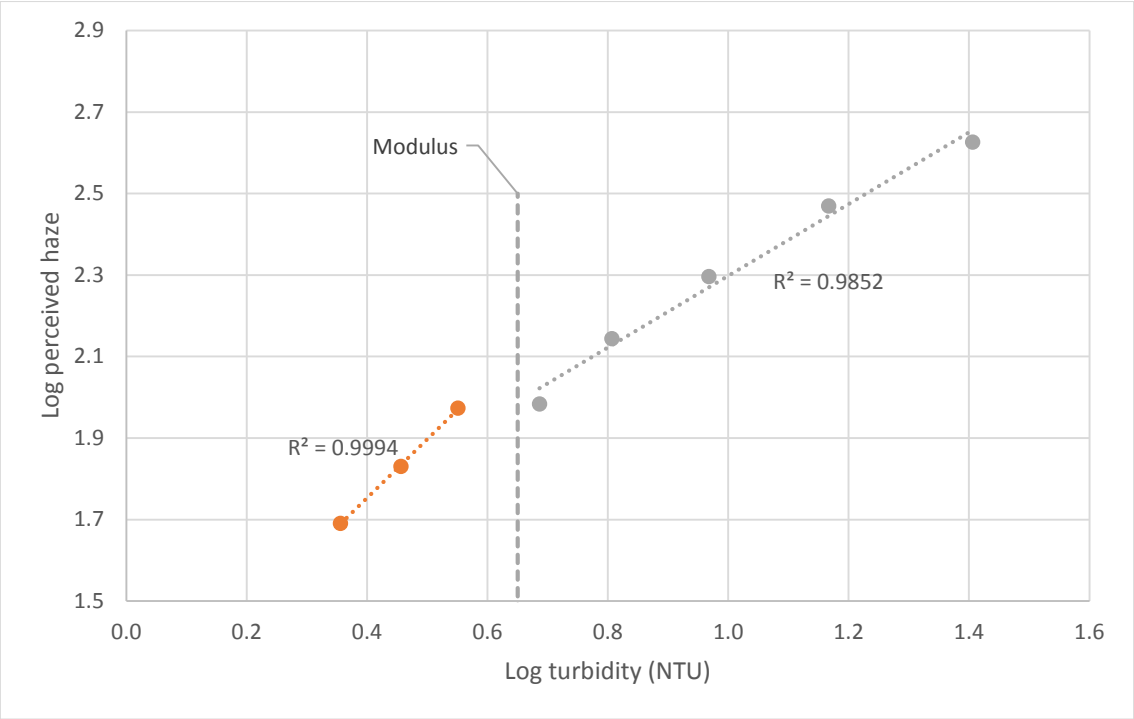


**Figure 6.3** **Log perceived haze versus log turbidity for haze induced Chardonnay wine samples. Straight line, linear best fits shown for upper and lower segments.**

# Chapter 7

# Comparison of Traditional Nephelometric Instrumental Measurement and Image Based Measurement of Polymer Microsphere Solutions

## 7.1 Overview and objectives

The overall goal of this study was to investigate the relationship between instrumental turbidity measurements and imaging data for yellow coloured samples containing polymer microspheres of a single nominal diameter (0.25 μm). The primary objective was to test the equivalence of measurements from a traditional nephelometer (Hach 2100P) with those predicted by a calibrated imaging model over a low turbidity range (< 6 NTU).

The imaging data was to be captured using of a modified Canon A2300 digital still camera along with a Samsung S2 smartphone device.

## 7.2 Materials and methods

The information given below should be read in conjunction with the general methodology Section found in Chapter 3.

### 7.2.1 Imaging devices

A critical feature of any camera device to be used to capture luminance data is access to unmodified linear sensor data. In preparation for this study the native features of a Samsung S2 smartphone were investigated and a number of limitations were found the most critical being that there was no facility to save images in a RAW sensor file format. The alternative options for the capture of sensor data were limited to JPEG file analysis or preview buffer extraction. Neither option provided access to unmodified linear sensor data.

A number of approaches were attempted to recover linear sensor data from non-linear JPEG and preview data including the use of the transformation algorithm described by Debevec and Malik[125]. Use of this method was unsuccessful due to the lack of manual exposure settings on the Samsung S2. Other alternatives such as applying a standard sRGB transformation as recommended by Szeliski[126] also failed due to the dynamic nature of the device's internal processing algorithms. Time limitations precluded further efforts forcing the exclusion of the Samsung S2 device from the study.

As previously described (Section 3.3.2) modifications were made to a Canon A2300 camera so that linear data could be extracted via raw image files. The Canon A2300 was the only camera device used in the study.

### 7.2.2 Sample preparation

A high turbidity yellow microsphere stock used in a previous experiment (Stock 10, 330 NTU, Table 5.2) was employed as a concentrated base to derive a series of 21 low turbidity samples through dilution with HPLC grade deionised water containing SDS and yellow pigment (at rates previously described in Section 3.4.2). The concentrated stock had been prepared approximately 4 months prior to this study (stored at 4˚C). Before using the concentrated stock the container was gently rotated 10 times to assist in microsphere dispersal. The range was designed to extend from the lowest turbidity possible up to around 6.0 NTU; within this range the concentration used for each sample was randomly selected (details in Table 7.1).

**Table 7.1 Details of yellow microsphere suspensions.**

| Sample | Microsphere density (number per mL) |
|---|---|
| 1 | $4.42 \times 10^7$ |
| 2 | $2.25 \times 10^7$ |
| 3 | $7.36 \times 10^7$ |
| 4 | $3.50 \times 10^6$ |
| 5 | $1.56 \times 10^7$ |
| 6 | $4.99 \times 10^6$ |
| 7 | $4.12 \times 10^7$ |
| 8 | $1.43 \times 10^7$ |
| 9 | $4.30 \times 10^7$ |
| 10 | $6.61 \times 10^7$ |
| 11 | $5.22 \times 10^7$ |

| Sample | Microsphere density (number per mL) |
|---|---|
| 12 | $3.65 \times 10^7$ |
| 13 | $1.54 \times 10^7$ |
| 14 | $2.21 \times 10^7$ |
| 15 | $6.74 \times 10^7$ |
| 16 | $4.29 \times 10^7$ |
| 17 | $9.14 \times 10^6$ |
| 18 | $1.80 \times 10^7$ |
| 19 | $6.40 \times 10^7$ |
| 20 | $5.20 \times 10^6$ |
| 21 | $8.19 \times 10^7$ |
| | |

### 7.2.3 Sample assessment

#### Anticipated interferences

In a range of pilot experiments using 90˚ overhead illumination it was found that the choice of glass receptacle had little impact on the linearity of imaging data when the turbidity of the sample was above 5.0 NTU. At lower turbidities it was found that reflections and liquid lensing became significant interfering factors. The observed effects included internal reflections, liquid surface reflections, external glass reflections and lensing of surround inhomogeneities. The imaging scene and receptacle were designed to minimise these effects but it is reasonable to expect some level of impact with lowest turbidity samples likely to be affected the most.

### Imaging scene

A small cylindrical jar made of consumer food grade glass of unknown optical characteristics (diameter: 3 cm; height: 12 cm) was used to hold the samples during image capture. The upper part of the cylinder (top 7.5 cm) was covered in opaque tape as a means on minimising reflections from the sample surface. The imaging cylinder containing the sample was located centrally in a cardboard imaging box (height: 14 cm, width: 22.5 cm, depth 14 cm). The imaging box had a rectangular port (18 cm x 10.5 cm) cut in the top surface to allow direct imaging of the sample. A second circular port (11 cm diameter) was cut in the front piece to allow placement of the sample. The imaging box was illuminated horizontally by two domestic grade 40 W tungsten filament bulbs. Each end of the image box had an open circular port (11 cm) to allow the bulbs to illuminate the sample; simple paper diffusers were located between the bulbs and the sample area. The base of the image box was covered with dark coloured paper overprinted with a template that allowed the imaging cylinder to be consistently positioned and aligned in a central location. The Canon A2300 camera was positioned vertically over the imaging cylinder (10 cm between the camera case and the top of the cylinder) with an unobscured view of the sample surface. The sample was not subject to any other illumination. The experimental setup is shown in Figure 7.1.

### Image capture and processing

Pilot experimentation showed that an exposure time of 500 ms was short enough to prevent saturation in the RGB channel responses. This exposure time was used for all images. Optical zoom was set to its zero setting; all other camera parameters were as described in A.3.

Processing of the images to arrive at RGB channel values were conducted in the manner previously described with a circular, central region of interest (ROI) used corresponding to a bounding rectangle of 250x250 pixels. The image data was split into its RGB channels and the green channel values were used as a basis for analysis.

### Measurement sequence

A graduated cylinder was used to pour 50 mL of each sample into the imaging cylinder. The imaging cylinder was then positioned on the template and the image captured. The sample was then poured into the Hach 2100P nephelometer cell (15 mL) from the imaging cylinder and the average turbidity was then measured. All glassware was flushed with sample material prior to use. The order in which the samples were imaged and turbidity measured was randomised (see Table 7.2). Each sample was imaged/measured once.

### 7.2.4 Statistical analysis

The parametric paired t-test was calculated using Microsoft Excel 2013's Data Analysis tool. The non-parametric Wilcoxon's Signed Ranks Test was manually computed in Microsoft Excel 2013 using critical values from O'Mahony[115].

## 7.3 Results and discussion

Green channel imaging values and turbidity measurements are listed in Table 7.2. As an initial check for method equivalence the turbidity values were plotted versus the corresponding green channel values; Figure 7.2. It can be seen that a reasonable linearity exists over the test range with an obvious group of three outlier points reducing the overall coefficient of determination ($R^2$=0.9628).

**Table 7.2 Imaging and instrumental turbidity data (results in order of measurement).**

| Sample | Canon A2300 green channel value[1] | Average turbidity measured by Hach 2100P (NTU) |
|---|---|---|
| 21 | 2.36 | 5.76 |
| 7 | 1.63 | 3.25 |
| 2 | 1.44 | 1.30 |
| 14 | 1.27 | 1.07 |
| 11 | 1.76 | 3.83 |
| 6 | 0.99 | 0.75 |
| 20 | 0.97 | 0.73 |
| 1 | 1.69 | 3.26 |
| 12 | 1.50 | 2.68 |
| 19 | 1.91 | 4.00 |
| 17 | 1.00 | 0.83 |
| 4 | 0.96 | 0.59 |
| 15 | 2.06 | 4.71 |
| 16 | 1.69 | 3.17 |
| 8 | 1.32 | 1.33 |
| 10 | 2.03 | 4.62 |
| 9 | 1.64 | 3.23 |
| 13 | 1.12 | 1.26 |
| 18 | 1.18 | 1.46 |
| 3 | 2.20 | 5.38 |
| 5 | 1.20 | 1.37 |

[1] For calculation convenience the raw green channel values had been divided by a factor of 1000 prior to inclusion in this table.

### 7.3.1 Outlier investigation

The three images corresponding to the outlier samples were carefully examined and it was found that two of the images contained a large number of bright suspended particulates likely causing elevated green channel values. This effect can be seen in Figure 7.3 where Sample 2 and Sample 13

are compared. These samples have similar instrumental turbidity values but significantly different green channel values. The visual differences are clear with the image of Sample 2 containing a large number of bright spots in comparison to Sample 13. A similar level of bright spots were found in a second outlier image: Sample 8. Visual examination of the third outlier image (Sample 14) did not show unusually large numbers of particulates. It was decided to exclude all three the outliers prior to statistical analysis resulting in an increased $R^2$ value of 0.9965.

### 7.3.2 Equivalence analysis

The analysis of equivalence is based on the premise that any practical implementation of an image based turbidity measurement system would probably use a pair of samples of known turbidity to derive a linear model for the prevailing scene. To test the viability of such an approach, it was first necessary to select two calibration points from the experimental data. Prior to conducting the experiment it had been decided that the upper calibration point should correspond to the highest turbidity sample as measured by the Hach 2100P nephelometer; this corresponded to Sample 21. The lower calibration point was to be chosen in the region of 1.5 NTU; this corresponded to Sample 18. These choices were somewhat arbitrary but the lower calibration point was chosen to avoid the sub 1 NTU area known to exhibit high nephelometric measurement variance[28,127]. Using these calibration points a linear model was derived as shown in Equation 7.1.

$$Predicted\ turbidity = 3.64 \times Green\ channel\ value - 2.82 \qquad 7.1$$

Equation 7.1 was then used to calculate a set of predicted turbidity values for each of the remaining 16 data points; the results are shown in Table 7.3. It can be seen that the differences between measured and predicted values appear relatively small for the most part.

**Table 7.3 Turbidity values as measured by Hach 2100P nephelometer and predicted turbidity values derived from Equation 7.1 based on experimental green channel data.**

| Sample | Measured turbidity Hach 2100P (NTU) | Predicted turbidity derived from green channel values (NTU) | Differences (NTU) |
|---|---|---|---|
| 1 | 3.26 | 3.32 | -0.06 |
| 3 | 5.38 | 5.20 | 0.18 |
| 4 | 0.59 | 0.68 | -0.09 |
| 5 | 1.37 | 1.53 | -0.16 |
| 6 | 0.75 | 0.77 | -0.02 |
| 7 | 3.25 | 3.12 | 0.13 |
| 9 | 3.23 | 3.15 | 0.08 |
| 10 | 4.62 | 4.58 | 0.04 |
| 11 | 3.83 | 3.59 | 0.24 |
| 12 | 2.68 | 2.63 | 0.05 |
| 13 | 1.26 | 1.24 | 0.02 |
| 15 | 4.71 | 4.66 | 0.05 |
| 16 | 3.17 | 3.32 | -0.15 |
| 17 | 0.83 | 0.82 | 0.01 |
| 19 | 4.00 | 4.12 | -0.12 |
| 20 | 0.73 | 0.70 | 0.03 |

A visual assessment of equivalence can be made by examining the scatter plot of predicted turbidity versus measured turbidity; this is shown in Figure 7.4. It can be seen that the data points are closely aligned to the 45˚ line providing superficial support for equivalence. To test statistical equivalence of the model predictions and instrumental measurements, the following hypotheses were proposed:

$H_0$: mean of differences = 0

$H_a$: mean of differences ≠ 0

A parametric paired two sample test for means resulted in a p-value of 0.61 (pre-set level of significance: 0.05). The null hypothesis was therefore accepted providing support for equivalence. Since the sample size does not allow a clear determination of normality of the distribution of mean differences, a non-parametric Wilcoxon Signed Ranks Test was also performed using a median difference of zero as the null hypothesis. This also resulted in acceptance of the null hypothesis (W=59, lower critical value =29, upper critical value =107) providing further evidence that the imaging model is reasonably equivalent to the Hach 2100P measurements for the given experimental conditions.

## 7.4 Summary and conclusions

The results from this study show that image based turbidity measurements may be a viable replacement for traditional instrumentation for homogenous, monodisperse samples over low NTU ranges. Where suspended particulates are present, image based measurements may deviate significantly from those provided by standard nephelometers. This effect is likely caused by the relatively large sample volumes and broad illumination patterns used for imaging leading to accumulation of bright spot artefacts. These are novel results and no other relevant published data is available for comparison.

The importance of the calibration points to be used in any practical system should be noted. The use of a single measurement for establishing calibration value could have a significant impact on the derived linear model. For example, the anomalous outlier in this study that had no unusual particulate presence may have been used for calibration leading to a significantly different outcome in the equivalence testing. The design of an image based turbidity measurement system would need to carefully consider the choice of calibration material to ensure stability. In addition, the measurement process would require replication and averaging of calibrant image data prior to model derivation.

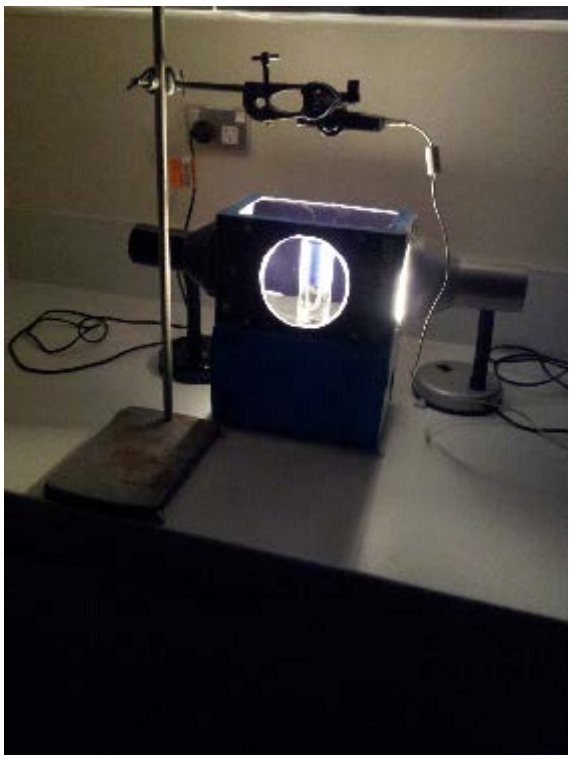

**Figure 7.1 Experimental setup for image capture.**

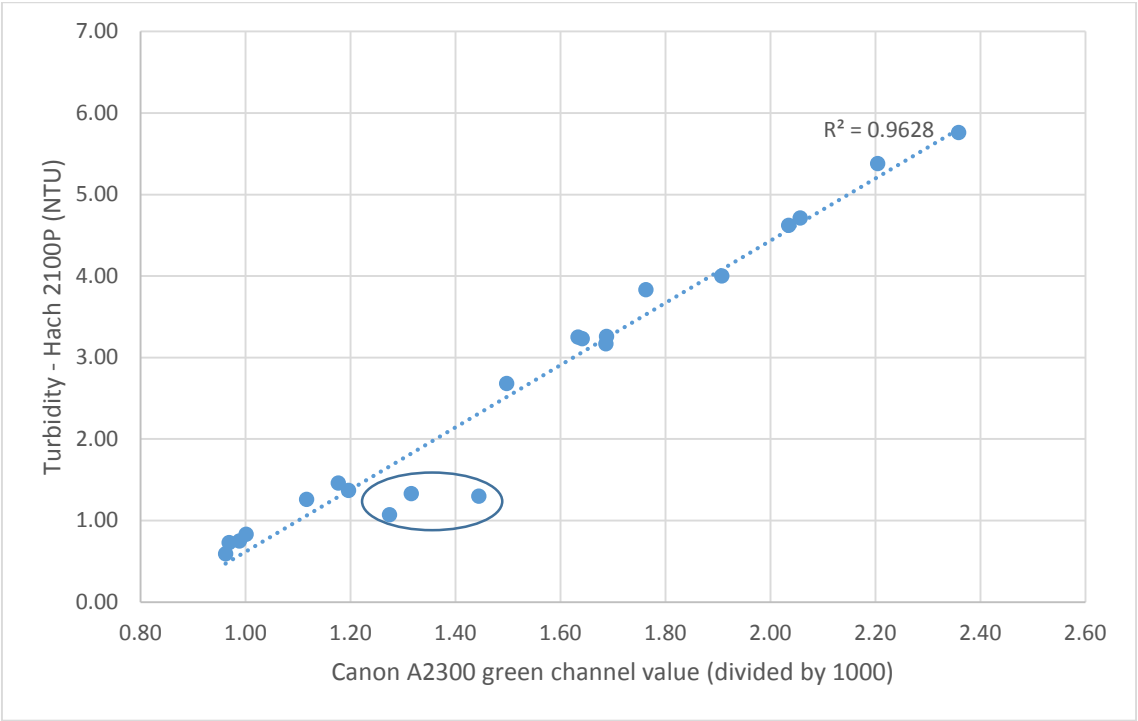


**Figure 7.2 Hach 2100P instrumental turbidity versus Canon A2300 green channel values. Best fit straight line regression shown ($R^2$=0.9628). Outlier group circled.**

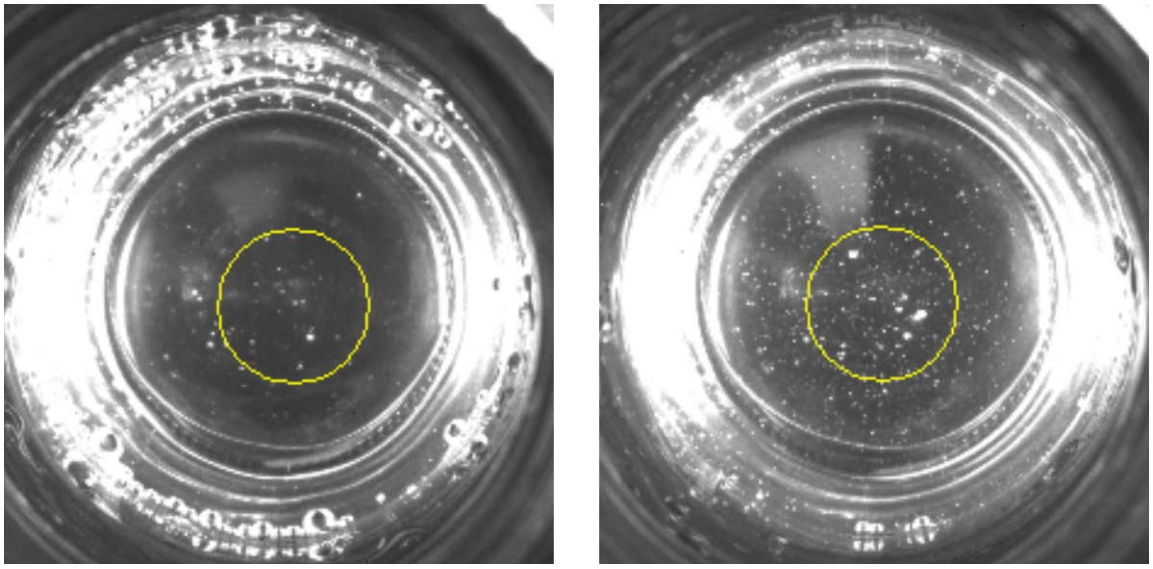


**Figure 7.3** **Two samples with similar NTU values but substantial differences in green channel values. Elevation of green channel values likely caused by suspended particulates.**

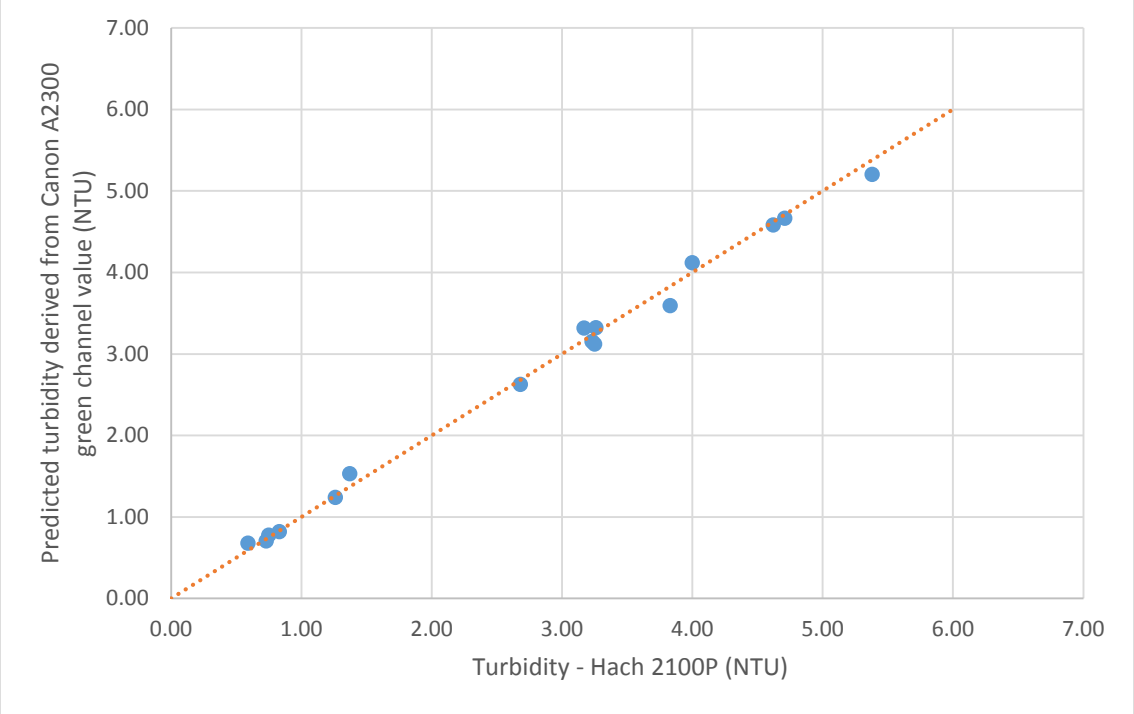


**Figure 7.4** **Predicted turbidity values derived from green channel data using Equation 7.1 versus Hach 2100P turbidity measurements. 45˚ line of equivalence included to illustrate close agreement of predicted and measured values.**

# Chapter 8

# Summary and Future Work

## 8.1 Haze detection thresholds in polymer microsphere suspensions and Chardonnay wine samples

In their study of haze detection thresholds in monodisperse microsphere suspensions Fleet and Siebert[6] found that human subjects could detect haze at turbidity levels of 0.2 NTU in achromatic suspensions and 0.43 NTU in tan coloured suspensions when evaluated in a brightly lit viewing box with a dark surround. They also found that changing from a dark to a light surround substantially increased the observed thresholds (12.8 NTU and 19.8 NTU respectively). This study has extended the work of Fleet and Siebert by employing conditions that more closely resemble those found when beverages are consumed in the real world. Haze detection thresholds in yellow microsphere suspensions presented in a standard tasting glass and viewed against a dark surround were found to be 1.15 NTU. On changing to a light surround the threshold was found to increase to 6.93 NTU. A Wilcoxon Signed Rank test indicated that the surround type was significant for both achromatic and yellow suspensions. Direct comparison of these results and those of Fleet and Siebert should be treated with caution due to possible inconsistencies in nephelometric measurements (see Section 2.2.2). In general terms the results are in broad agreement with a possible tempering of the extremes by real-world conditions where higher thresholds were found for the dark surround and lower thresholds for the light surround.

Wine has a complex composition and extrapolating haze detection levels in monodisperse microsphere suspensions to wine may not be appropriate. In this study thresholds in heat treated Chardonnay wine samples were found to be 0.86 NTU in a dark surround. Although this is somewhat lower than the equivalent 1.15 NTU found for the yellow microsphere suspensions, a Wilcoxon Signed Rank test did not detect a significant difference.

## 8.2 Measuring visual haze

If the visual estimation of haze in wine is to be replaced by an equivalent instrumental technique, it is first necessary to obtain a mathematical model that can transform luminance measurements into the perceptual domain. To this end Magnitude Estimation[128] and a modified Rank-Rating[118] method were used to scale the visual response of human subjects to monodisperse polymer microsphere suspensions and heat treated Chardonnay wine samples. As with the threshold study, the experimental conditions employed resembled those found when beverages are consumed in the real-world with both light and dark surrounds. A number of generalised mathematical functions

derived from empirical psycho-physical models of brightness/lightness perception were compared to the experimental data. Although reasonable fits were found for both logarithmic and power forms, individual parameter values were necessary for each suspension colour and surround. It is not clear how these parameter values might be practically derived to allow the human response to be estimated for some specific hazy wine sample located in a real-world scene composed of other inducting elements.

A key observation from this work was the likely inductive interaction between the test sample and other scene elements (such as a modulus in Magnitude Estimation). This means that any model that only deals with gross surround induction without considering other scene element attributes (e.g. luminance, relative area and planar positioning) is likely to fail[74,129].

## 8.3 Nephelometry and digital imaging

The relationship between nephelometric measurements and imaging data was investigated for monodisperse polymer microsphere suspensions presented in standard sampling glasses located in a booth used for sensory experimentation (90˚ lighting). It was found that data derived from digital photographs of the samples in a dark surround had a strong linear relationship with nephelometric measures below 50 NTU ($R^2$=0.9978). However, at the lower turbidity ranges (< 5 NTU), a number of interfering effects such as internal reflections, liquid surface reflections, external glass reflections and lensing of surround inhomogeneities caused inconsistencies in the image data.

Using a different photographic configuration where the surface of heat treated wine samples were imaged directly, it was possible to obtain consistent results over a low range (0.59 NTU to 5.76 NTU). Predicted turbidity values based on imaging data were compared with nephelometric measurements using a parametric and non-parametric tests; no significant differences between the results were found. Although this suggests that image based nephelometry may be a viable alternative to traditional nephelometric instruments, a small number of outliers were found in the data (excluded from the statistical tests). In some of the relevant images larger particulates were visible causing bright spots that biased the image data. Traditional nephelometers are restricted to narrow fields of view and a small number of discrete viewing angles. In contrast, a digital image of a sample can be used to derive information about the entire volumetric sample. These measurement system differences are likely to cause differences in observed data especially when the sample is non-homogenous.

### 8.3.1 Future work

The derivation of a suitable model for transforming luminance measurements into perceptual haze intensities will require a more sophisticated approach than the simple logarithmic and power

functions explored in this study. Such a model will need to accommodate the interaction of all elements in the field of view not just the liquid sample. It will also need to recognise chromatic interactions such as the Helmholtz–Kohlrausch Effect. Progress in this area is likely to require the use of a full colorimetric appearance model specifically designed to accommodate inductive effects; CIECAM-m2[101] is one such candidate. Development in tone mapping algorithms using Anchoring Theory may also provide some useful insights[130,131]. Successful assessment of low haze samples is also likely to require the use of a suitable model for real-world, three dimensional transparency.

Although the potential for image based turbidity measurement has been demonstrated, the development of a useful system will require further design and testing. Imaging geometry, illumination, receptacle optics, sample volume, calibration material and stray light management will all need to be optimised through engineering prototype studies. In addition, the use of formal Measurement System Analysis techniques (such as Gauge Repeatability and Reproducibility studies)[132] is recommended along with statistical method equivalence testing[133,134].

One of the limiting factors in the imaging aspects of this study has been the difficulty in extracting unmodified linear sensor data from smartphone devices. This issue is well recognised and work is currently underway to improve the computational camera functionality in mobile platforms through developments such as Android HAL3[135]. When these enhancements become available, the prospects for nephelometric measurement using a mobile device will improve significantly.

# Appendix A

# Image Capture and Processing Pipeline

## A.1 Overview

Image capture and associated data processing employed in this study is outlined in Figure 8.1. Key features and parameter values for each pipeline element are detailed below.

## A.2 Canon A2300

The Canon A2300 is a standard consumer quality digital still camera with a charge coupled device (CCD) sensor consisting of 16 mega-pixels. The camera does not have a manual mode and can only save images in the compressed JPEG format.

The A2300 was loaded with a non-Canon customised firmware known as the Canon Hack Development Kit (CHDK)[136]. CHDK is an open source software platform that provides an application programmatic interface allowing script based interaction with a range of standard and non-standard features in Canon cameras. Importantly, this includes the facility set exposure times and to save image files as Digital Negative Graphic (DNG)[137] files. DNG is a lossless format that provides an open standard for exchange and storage of raw image files independent of proprietary raw file formats used by camera manufacturers. In the case of the A2300 camera, the DNG file output contains unmodified colour filter array data with a resolution of 12 bits per pixel.

Pilot work to assess the response of the end to end system indicated that the image processing pipeline was capable of providing linear device RGB data if the exposure was set at a level that avoided channel saturation (see Appendix B).

## A.3 Image capture

Image capture was controlled using customised Lua[138] scripts executed via a USB interface utility: PTPCamGui[139]. This approach allowed programmatic control of the exposure as well as providing a means of automatically logging the details of each image exposure.

The vast majority of the camera parameters were maintained as constants throughout the study with exposure time being varied to ensure that sensor saturation was avoided. For each experimental configuration this was achieved by taking trial images of the most luminous sample in the sample set and selecting an exposure time that would avoid saturated sensor responses. Once an appropriate exposure time was chosen it was fixed for each experimental scene (as was the camera zoom level).

It is known that the amount of light illuminating the sensor array in a camera varies; this effect increases with radial angle from the central optical axis[140]. To minimise these vignetting effects the scene setup was designed to maintain the target sample Region of Interest (ROI) within the central 50% of the image area.

Prior to experimental work commencing the camera was calibrated for bad pixels using the CHDK standard procedure. This ensured that defective pixels values were suitably interpolated prior to saving in the DNG format.

A list of the key camera parameters and values set are shown in the table below; all other parameters were maintained as default values.

| Parameter | Value |
|---|---|
| DNG Version | 1.3 |
| Neutral density filter | “out” |
| ISO | 100 |
| Autofocus | enabled |

## A.4 DNG to PGM conversion

The DNG files were converted to Portable Grey Map (PGM) format using the DCRAW conversion utility[141]. The command line argument string (“-4 –D”) was used to prevent alteration of the raw pixel values through Bayer demosaicing processes or through the application of a colour profile.

## A.5 Derivation of channel pixel values

Final derivation of device RGB data (12 bits/pixel) for selected regions of interest (ROI) in image files was accomplished through the use of the public domain software IMAGEJ[142]. A plugin sourced from the University of Manitoba[143] was used to “debayer” the image data into individual channels ("order=G-R-G-R, demosaicing=Replication, radius=2, radius=2). The device RGB values in the region of interest were then used to calculate the aggregate luminance values using Equation 3.1.

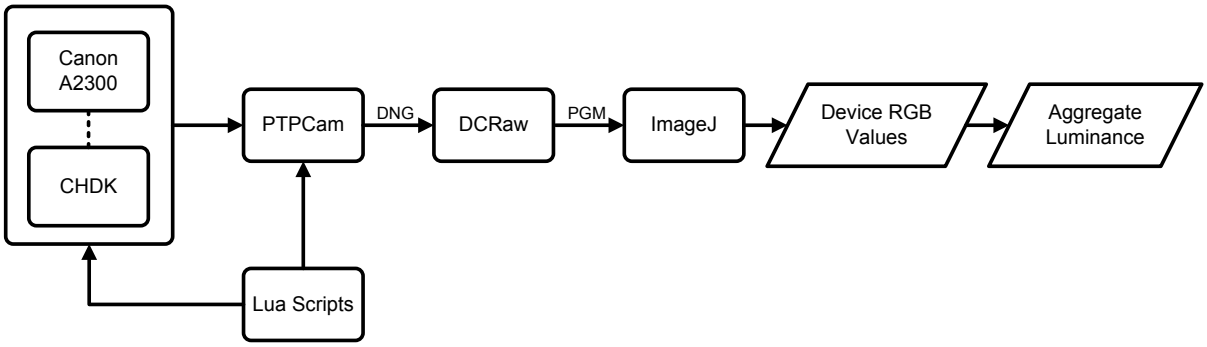


**Figure 8.1 Image capture and processing pipeline employed in this study.**

# Appendix B

# Image Capture and Processing Linearity Tests

## B.1 Overview and objectives

The Canon A2300 digital still camera uses a charge coupled device (CCD) array to collect information from light delivered to it by its front end optics. It is generally known that the charge collected by a CCD is proportional to its irradiance[125]. Although the Canon Hack Development Kit (CHDK) modifications made to the camera should ensure that linear sensor data is present in the raw DNG file output from the device, it was thought prudent to test the linearity of the modified camera in conjunction with the overall image processing pipeline.

## B.2 Methods and materials

### B.2.1 Imaging setup

A 1 cm square black outline printed on white photocopier quality paper was used as a target. Illumination was provided by a domestic grade angle-poise lamp containing a 40 W incandescent bulb. The lamp was fixed in place so that it could not move and switched on 30 minutes prior to taking the first image. The lamp illuminated the target at an angle approximating 45˚ at a distance of 40 cm. The target was not illuminated by any other source.

### B.2.2 Image capture

The camera was directly aligned (orthogonal) to the target at a distance of 10 cm. Focus was set to zero; all other camera parameters were as described in A.3. A total 96 images were captured using 32 discrete exposure settings ranging from 1 ms to 1000 ms (each discrete exposure was used in three images). The order in which the exposures were applied was randomised.

Processing of the images to arrive at RGB channel values were conducted in the manner previously described (see Section 3.3.2) using a rectangular region of interest (ROI) centred on and inside the rectangular target corresponding to 400x400 pixels. The image data was split into its RGB channels in preparation for analysis.

## B.3 Results

The channel values for each of the exposed images are shown Figure B.1 up to an exposure of 250 ms. It can be seen that the red and green channels have a similar sensitivity to the light reflected from the target with the blue channel somewhat less sensitive. Each channel shows a similar

characteristic: a linear segment followed by a flat segment representing saturation in the sensor (i.e. the maximum value available from the 12 bit channels: 4095). There is insufficient data points to be certain but linearity may be lost below channel values of 4000.

Figure B.2 shows the data truncated at a maximum of 3500 from each of the channels. It can be seen that each channel is strongly linear (red $R^2$=0.9999, green $R^2$=0.999, blue $R^2$=0.9997).

## B.4 Conclusions

It was concluded that the modified Canon A2300 camera in conjunction with the processing pipeline defined in Appendix A would reliably deliver linear data. To minimise the risk of entering saturation exposure times in all experiments were to be designed to ensure that relevant channel values remained below 3000.

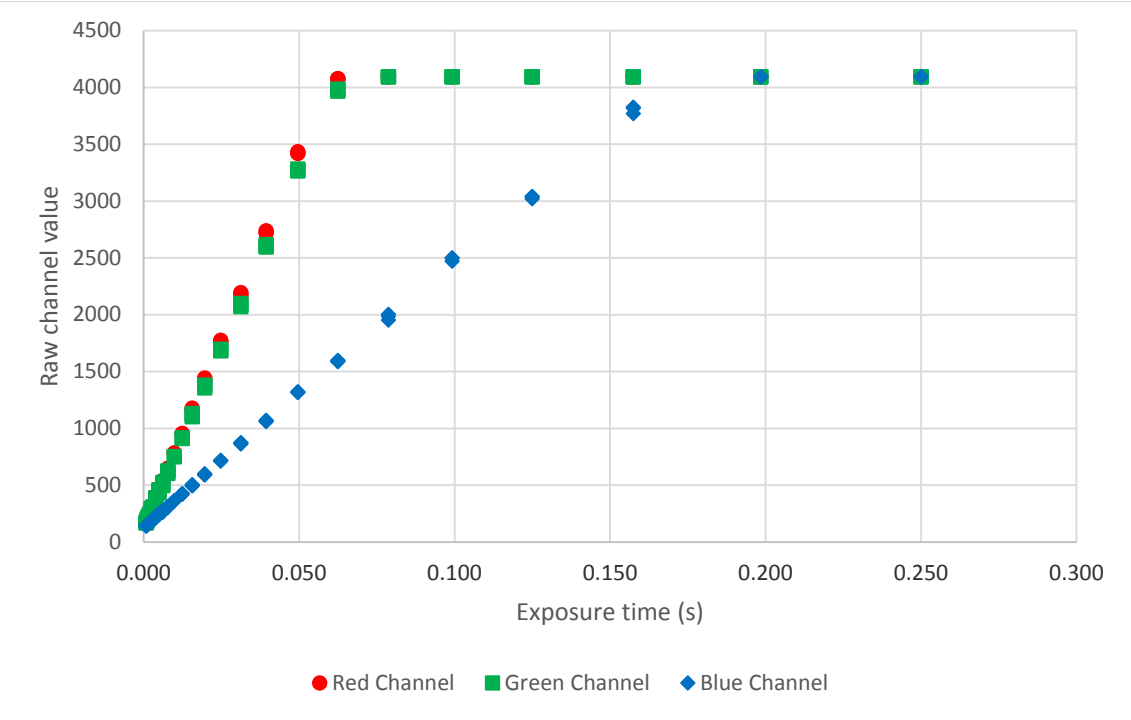


**Figure B.1 Channel values at the output of the processing pipeline based on a range of exposure times used to capture an image of a fixed target. Three images were taken at each of the 32 selected exposure times (most data points are closely superimposed). Image data for exposures above 250 ms are all in saturation and not shown.**

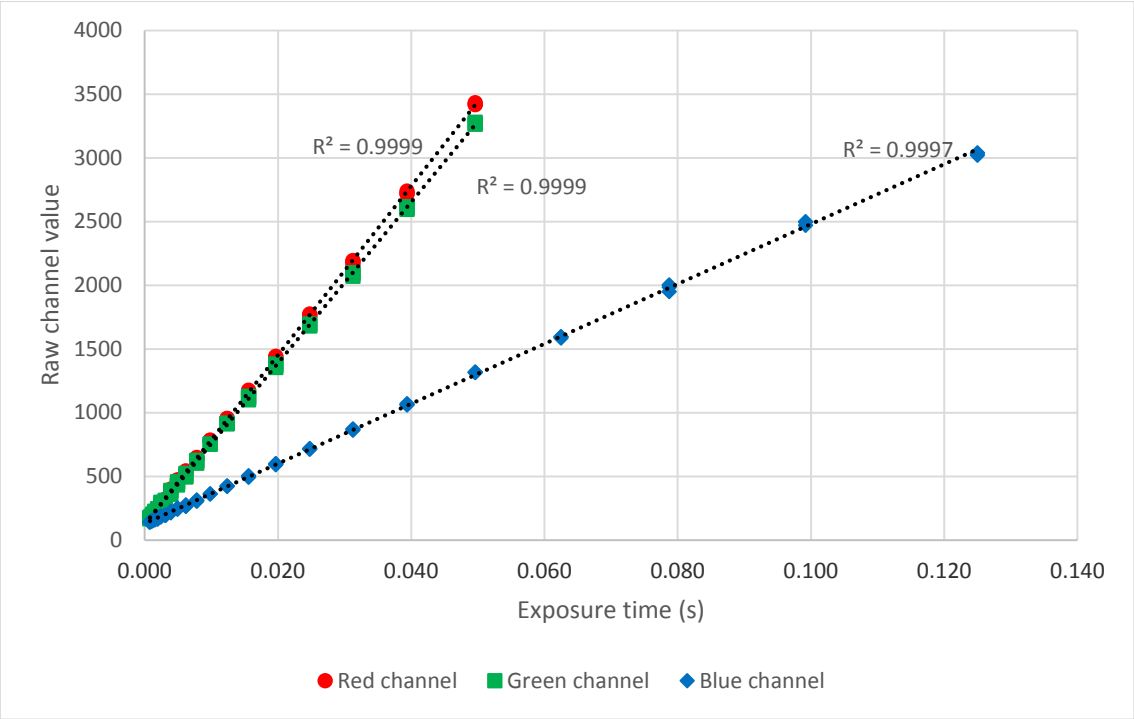


**Figure B.1 Channel values at the output of the processing pipeline. Results truncated at a maximum channel value of 3500.**

# Appendix C

# Pre-written Scripts Used For Briefing Subjects In Sensory Experiments

## C.1 Subject instruction script - haze detection threshold experiment using polymer microsphere suspensions

*In this experiment you will be evaluating liquid samples presented in standard wine glasses. The samples contain polymer beads and should not be consumed.*

*The experiment will commence with the presentation of a set of three glasses and your job is to evaluate the samples and decide which is the haziest. The glasses will be labelled with a three digit label; you should record your decision on the results sheet by circling the appropriate code. You will be given 30 seconds to make your choice at which point you will hear an audible alarm. If you can't decide which is the haziest just guess. When you hear the alarm the hatch will be opened and the glasses will be replaced. This will continue until you have evaluated ten sets of glasses. At this point I will return to the sensory room and direct you to the next sensory booth; you will evaluate ten sets of samples in four different booths.*

*When you are presented with a set of glasses to be evaluated, you should compare each of the three possible pairs of glasses by raising them to eye level. When you are doing this please hold the glasses by the stem. Once this is completed you are free to manipulate and examine the glasses in any way you feel appropriate but you should avoid assessing a sample over the score sheet; try to keep the score sheet off to one side.*

*I will remain silent when I'm presenting the samples. If you have any problems during the course of the evaluations just knock on the hatch. If there is a spill at any point, we will stop the evaluations, clean up before restarting. If you are presented with samples that do not match the sample set on your score sheet please knock on the hatch.*

## C.2 Subject instruction script - haze detection threshold experiment using Chardonnay wine samples

*In this experiment you will be evaluating wine samples presented in standard wine glasses; the samples should not be consumed.*

*The experiment will commence with the presentation of a set of three glasses and your job is to evaluate the samples and decide which is the haziest. The glasses will be labelled with a three digit*

*label; you should record your decision on the results sheet by circling the appropriate code. You will be given 30 seconds to make your choice at which point you will hear an audible alarm. If you can't decide which is the haziest just guess. When you hear the alarm the hatch will be opened and the glasses will be replaced. This will continue until you have evaluated ten sets of glasses.*

*When you are presented with a set of glasses to be evaluated, you should compare each of the three possible pairs of glasses by raising them to eye level. When you are doing this please hold the glasses by the stem. Once this is completed you are free to manipulate and examine the glasses in any way you feel appropriate but you should avoid assessing a sample over the score sheet; try to keep the score sheet off to one side.*

*I will remain silent when I'm presenting the samples. If you have any problems during the course of the evaluations just knock on the hatch. If there is a spill at any point, we will stop the evaluations and clean up before restarting. If you are presented with samples that do not match the sample set on your score sheet please knock on the hatch.*

## C.3 Subject instruction script - haze intensity scaling experiment using polymer microsphere suspensions

*In this experiment you will be evaluating the intensity of haze in a range of liquid samples presented in standard wine glasses. The samples contain polymer beads and should not be consumed.*

*The experiment will commence with the presentation of a glass containing the first test sample. At this point you should compare the test sample to the reference sample marked with an "X". You are to judge the comparative haziness just as you did in the training session by rating the test sample in proportion to the reference sample where the reference sample has an assigned value of 100. For example, if the test sample is twice as hazy as the reference, assign it a value of 200. If the test sample is one tenth as hazy as the reference, assign it a value of 10. You can use any positive numbers to assign values to the test sample. Where you can't detect any haze at all in the test sample, assign it the value of zero.*

*The glasses will be labelled with a three digit label; you should record your decision on the results sheet alongside the relevant code. You will be given 30 seconds to make your evaluation at which point you will hear an audible alarm. When you hear the alarm the hatch will be opened and the test sample will be replaced. This will continue until you have evaluated eight test samples. At this point I will return to the sensory room and direct you to the next sensory booth; you will evaluate eight samples in four different booths.*

*When you are presented with a test sample to be evaluated, you should compare the test sample with reference by raising them both to eye level first. The reference glass should be held in your [left|right] hand and the test sample in your [left|right] hand. When you are doing this please hold the glasses by the stem. Once this is completed you are free to manipulate and examine the glasses in any way you feel appropriate but you should avoid assessing a sample over the score sheet; try to keep the score sheet off to one side.*

*I will remain silent when I'm presenting the samples. If you have any problems during the course of the evaluations just knock on the hatch. If there is a spill at any point, we will stop the evaluations and clean up before restarting. If you are presented with a sample that do not match the sample code on your score sheet please knock on the hatch.*

## C.4 Subject instruction script - haze intensity scaling experiment using Chardonnay wine samples

*In this experiment you will be evaluating the intensity of haze in a range of wine samples presented in standard wine glasses; they should not be consumed.*

*This experiment is a modified version of the experiment we conducted when scaling the intensity of haze in polymer bead samples. In this experiment you will be assessing the samples as a group over two stages. In the first stage your job will be to re-arrange a group of samples into a row where the least hazy sample is at the far [left|right] and the haziest is at the far [left|right]. In other words you should arrange them in order of haze [left to right | right to left]. When you are re-arranging the samples please hold the glasses by the stem.*

*Once you have arranged your samples in a row I will provide you with the reference sample marked with an "X" against which you should judge the test samples; you should hold the reference sample in your [left|right] hand. You are to judge the comparative haziness just as you did in the previous experiment by rating the test sample in proportion to the reference sample where the reference sample has an assigned value of 100. For example, if the test sample is twice as hazy as the reference, assign it a value of 200. If the test sample is one tenth as hazy as the reference, assign it a value of 10. You should start your evaluations at the left end of the row and continue rightwards until you have evaluated all samples. You can use any positive numbers to assign values to the test sample. Unlike the previous experiment you cannot evaluate a sample as zero; you must assign some non-zero value to each sample.*

*When judging a test sample, you should compare the test sample with reference by raising them both to eye level first. Again, please hold the glasses by the stem. Once you have completed the eye level*

*examination you are free to manipulate and examine the glasses in any way you feel appropriate. As your make your evaluations I will record your assigned scores.*

*At the end of the entire process you will be given the scores you have assigned and you will be free to re-evaluate any of the samples as you wish; you will be able to change any of the scores.*

*There will be no time limit imposed on your evaluations; just proceed through each stage at a rate you are comfortable with.*

*I will remain silent unless you have trouble with the process; you are free to ask for clarification at any point.*

# References


1. Iland, P. *Techniques for chemical analysis and quality monitoring during winemaking*. (Patrick Iland Wine Promotions, 2000).
2. Pocock, K. F. & Waters, E. J. Protein haze in bottled white wines: How well do stability tests and bentonite fining trials predict haze formation during storage and transport? *Aust. J. Grape Wine Res.* **12,** 212–220 (2006).
3. Carrasco, A. & Siebert, K. J. Human visual perception of haze and relationships with instrumental measurements of turbidity. Thresholds, magnitude estimation and sensory descriptive analysis of haze in model systems. *Food Qual. Prefer.* **10,** 421–436 (1999).
4. Horne, J., Olabi, A., Greenwalt, C. & Lawless, H. T. Visual Haze Detection Threshold Assessment by Ascending Method of Limits and a Transformed Staircase Procedure in Apple Juice Simulations Under ' Store Like' Conditions. *J. Sens. Stud.* **16,** 447–460 (2001).
5. Fleet, C. F. & Siebert, K. J. Effect of illumination intensity on visual perception of turbidity. *Food Qual. Prefer.* **16,** 536–544 (2005).
6. Fleet, C. F. & Siebert, K. J. Effect of Viewing Background on Turbidity Perception Thresholds. *J. Sens. Stud.* **21,** 34–53 (2006).
7. Majewski, P., Barbalet, A. & Waters, E. $1 billion hidden cost of bentonite fining. *Aust. N. Z. Grapegrow. Winemak.* 58–62 (2011). at <http://dialnet.unirioja.es/servlet/articulo?codigo=3684251>
8. New Zealand Winegrowers. Wineries by Region & Category April 2012. (2012).
9. Sauvage, F. X., Bach, B., Moutounet, M. & Vernhet, A. Proteins in white wines: Thermo-sensitivity and differential adsorbtion by bentonite. *Food Chem.* **118,** 26–34 (2010).
10. Sarmento, M. R., Oliveira, J. C., Slatner, M. & Boulton, R. B. Comparative quantitative analysis of the effect of cultivar, wine growing region and vinification method on the protein profiles of some white wines. *Int. J. Food Sci. Technol.* **36,** 759–766 (2001).
11. Monteiro, S. *et al.* The Wide Diversity of Structurally Similar Wine Proteins. *J. Agric. Food Chem.* **49,** 3999–4010 (2001).
12. Mesquita, P. R. *et al.* Effect of Wine Composition on Protein Stability. *Am. J. Enol. Vitic.* **52,** 324–330 (2001).
13. Dufrechou, M., Poncet-Legrand, C., Sauvage, F.-X. & Vernhet, A. Stability of White Wine Proteins: Combined Effect of pH, Ionic Strength, and Temperature on Their Aggregation. *J. Agric. Food Chem.* **60,** 1308–1319 (2012).
14. Waters, E. J. *et al.* Preventing protein haze in bottled white wine. *Aust. J. Grape Wine Res.* **11,** 215–225 (2005).
15. Dufrechou, M., Sauvage, F. X., Bach, B. & Vernhet, A. Protein Aggregation in White Wines: Influence of the Temperature on Aggregation Kinetics and Mechanisms. *J. Agric. Food Chem.* **58,** 10209–10218 (2010).
16. Pocock, K. F., Alexander, G. M., Hayasaka, Y., Jones, P. R. & Waters, E. J. Sulfate A Candidate for the Missing Essential Factor That Is Required for the Formation of Protein Haze in White Wine. *J. Agric. Food Chem.* **55,** 1799–1807 (2007).
17. Sauvage, F. X., Bach, B., Moutounet, M. & Vernhet, A. Proteins in white wines: Thermo-sensitivity and differential adsorbtion by bentonite. *Food Chem.* **118,** 26–34 (2010).
18. Weiss, W. F., Young, T. M. & Roberts, C. J. Principles, approaches, and challenges for predicting protein aggregation rates and shelf life. *J. Pharm. Sci.* **98,** 1246–1277 (2009).
19. Ribéreau-Gayon, P., Glories, Y., Maujean, A. & Dubordieu, D. *Handbook of Enology: The Chemistry of Wine and Stabilisation and Treatments, vol. 2 John Wiley & Sons Ltd*. (Chichester, 2000).
20. Gregory, J. Turbidity and beyond. *Filtr. Sep.* **35,** 63–67 (1998).
21. Kerker, M. *The Scattering Of Light*.
22. Thorne, R. S. W. & Nannestad, I. Some Considerations on the Physical Significance of Turbidity Estimates. *J. Inst. Brew.* **65,** 175–188 (1959).

23. Narasimhan, S. G. *et al.* Acquiring scattering properties of participating media by dilution. *ACM Trans Graph* **25,** 1003–1012 (2006).
24. Egan Jr, J. W. Turbidity–Clarifying Low Level Measurements. at <http://www.warws.com/pdfs/turbidity_article_rev_10.pdf>
25. Sadar, M. Making Sense of Turbidity Measurements–Advantages In Establishing Traceability Between Measurements and Technology. in *2004 Natl. Monit. Conf. Build. Sustain. Success. Monit. Programs Natl. Water Qual. Counc.* (2004). at <http://acwi.gov/monitoring/conference/2004/conference_agenda_links/papers/poster_papers/215_SadarMike.pdf>
26. Thorne, R. S. W. Application of Formazin Standards to Nephelometric Estimation of Beer Turbidity. *J. Inst. Brew.* **67,** 191–199 (1961).
27. Morris, T. M. The Relationship Between Haze and the Size of Particles in Beer. *J. Inst. Brew.* **93,** 13–17 (1987).
28. Letterman, R. D., AWWA Research Foundation & American Water Works Association. *A study of low-level turbidity measurements*. (AWWA Research Foundation and American Water Works Association, 2002).
29. Thorne, R. S. W. & Svendsen, K. Particle Size of Beer Turbidigens and Its Influence on Nephelometry. *J. Inst. Brew.* **68,** 257–270 (1962).
30. Francis, F. J. & Clydesdale, F. M. *Food colorimetry: theory and applications.* (AVI Publishing Co. Inc., 1975). at <http://www.cabdirect.org/abstracts/19760423502.html>
31. Zoecklein, B. W. *Production wine analysis*. (Van Nostrand Reinhold, 1990).
32. Gales, P. W. A comparison of visual turbidity with turbidity measured by commercially available instruments. *J. Am. Soc. Brew. Chem.* **58,** 101–107 (2000).
33. Hutchings, J. B. *Food color and appearance*. (Aspen Publishers Gaithersburg, 1999). at <http://www.getcited.org/pub/100418839>
34. Martin, M. L. G.-M., Ji, W., Luo, R., Hutchings, J. & Heredia, F. J. Measuring colour appearance of red wines. *Food Qual. Prefer.* **18,** 862–871 (2007).
35. Fleming, R. W. & Bülthoff, H. H. Low-Level Image Cues in the Perception of Translucent Materials. *ACM Trans Appl Percept* **2,** 346–382 (2005).
36. Gkioulekas, I. *et al.* Understanding the Role of Phase Function in Translucent Appearance. *ACM Trans Graph* **32,** 147:1–147:19 (2013).
37. Hunter, R. S. *The measurement of appearance*. (Wiley, 1987).
38. Nakauchi, S., Usui, S., Parkkinen, J. & Silfsten, P. Computational explanations for color transparency. in *Int. Jt. Conf. Neural Netw. 1999 IJCNN 99* **1,** 170–173 vol.1 (1999).
39. Metelli, F. The perception of transparency. *Sci. Am.* **230,** 90 (1974).
40. Metelli, F. Achromatic color conditions in the perception of transparency. (1974). at <http://psycnet.apa.org/psycinfo/1975-20146-015>
41. Metelli, F. Stimulation and perception of transparency. *Psychol. Res.* **47,** 185–202 (1985).
42. Singh, M. & Anderson, B. L. Toward a perceptual theory of transparency. *Psychol. Rev.* **109,** 492–518 (2002).
43. Singh, M. & Anderson, B. L. Photometric determinants of perceived transparency. *Vision Res.* **46,** 879–894 (2006).
44. Singh, M. Lightness constancy through transparency: internal consistency in layered surface representations. *Vision Res.* **44,** 1827–1842 (2004).
45. Vladusich, T. A reinterpretation of transparency perception in terms of gamut relativity. *JOSA A* **30,** 418–426 (2013).
46. Albert, M. K. Lightness and perceptual transparency. *Perception* **35,** 433 – 443 (2006).
47. Anderson, B. L. *The perceptual representation of transparency, lightness, and gloss*. (Handbook of perceptual organization. Oxford: Oxford University Press). at <http://www.gestaltrevision.be/pdfs/oxford/Anderson-The_perceptual_representation_of_transparency_lightness_and_gloss.pdf>
48. Beck, J. Perception of transparency in man and machine. in *Pap. Second Workshop Vol 13 Hum. Mach. Vis. II* 1–12 (1986). at <http://dl.acm.org/citation.cfm?id=39682>

49. Brill, M. H. Physical and informational constraints on the perception of transparency and translucency. *Comput. Vis. Graph. Image Process.* **28,** 356–362 (1984).
50. Faul, F. & Ekroll, V. On the filter approach to perceptual transparency. *J. Vis.* **11,** (2011).
51. Motoyoshi, I. Highlight–shading relationship as a cue for the perception of translucent and transparent materials. *J. Vis.* **10,** 6 (2010).
52. Nagai, T. *et al.* Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study. *iPerception* **4,** 407–428 (2013).
53. Malcolmson, L. J., Jeffery, L., Sharma, D. D. & Ng, P. K. W. Relationship Between Sensory Clarity and Turbidity Values of Apple Juice. *Can. Inst. Food Sci. Technol. J.* **22,** 129–132 (1989).
54. Carrasco, A. Human visual perception of haze and relationships with instrumental measurements of turbidity: Effects of polyphenol adsorbents and fining agents in grape juice and wine. (1997).
55. Wyszecki, G. & Stiles, W. S. *Color science: concepts and methods, quantitative data and formulae*. (Wiley, 1982).
56. McCann, J. J. The Appearance of Brightness and Lightness. *Color Imaging Conf.* **2000,** 18–23 (2000).
57. Vladusich, T. Brightness scaling according to gamut relativity. *Color Res. Appl.* (2013). doi:10.1002/col.21823
58. Blakeslee, B., Reetz, D. & McCourt, M. E. Coming to terms with lightness and brightness: Effects of stimulus configuration and instructions on brightness and lightness judgments. *J. Vis.* **8,** 3 (2008).
59. Withouck, M. *et al.* Brightness perception of unrelated self-luminous colors. *J. Opt. Soc. Am. A* **30,** 1248–1255 (2013).
60. Reinhard, E. *Color imaging: fundamentals and applications*. (A.K. Peters, 2008).
61. Kingdom, F. & Moulden, B. A model for contrast discrimination with incremental and decremental test patches. *Vision Res.* **31,** 851–858 (1991).
62. Whittle, P. in *Pigments Percept.* 293–304 (Springer, 1991). at <http://link.springer.com/chapter/10.1007/978-1-4615-3718-2_35>
63. Adams, E. Q. & Cobb, P. W. The Effect on Foveal Vision of Bright (and Dark) Surroundings. V. *J. Exp. Psychol.* **5,** 39–45 (1922).
64. Shapley, R. & Reid, R. C. Contrast and assimilation in the perception of brightness. *Proc. Natl. Acad. Sci.* **82,** 5983–5986 (1985).
65. Stevens, S. S. To Honor Fechner and Repeal His Law: A power function, not a log function, describes the operating characteristic of a sensory system. *Science* **133,** 80–86 (1961).
66. Jameson, D. & Hurvich, L. M. Perceived Color and Its Dependence on Focal, Surrounding, and Preceding Stimulus Variables. *JOSA* **49,** 809–897 (1959).
67. Bressan, P. Explaining lightness illusions. *Percept.-Lond.*- **30,** 1031–1046 (2001).
68. Takasaki, H. Lightness Change of Grays Induced by Change in Reflectance of Gray Background. *J Opt Soc Am* **56,** 504–509 (1966).
69. Whittle, P. Increments and decrements: Luminance discrimination. *Vision Res.* **26,** 1677–1691 (1986).
70. Whittle, P. Brightness, discriminability and the 'Crispening Effect'. *Vision Res.* **32,** 1493–1507 (1992).
71. Semmelroth, C. C. Prediction of lightness and brightness on different backgrounds. *JOSA* **60,** 1685–1689 (1970).
72. Gilchrist, A. L. & Radonjić, A. Anchoring of lightness values by relative luminance and relative area. *J. Vis.* **9,** 13 (2009).
73. Shi, V., Cui, J., Troncoso, X. G., Macknik, S. L. & Martinez-Conde, S. Effect of stimulus width on simultaneous contrast. *PeerJ* **1,** e146 (2013).
74. Radonjić, A. & Gilchrist, A. L. Depth effect on lightness revisited: The role of articulation, proximity and fields of illumination. *iPerception* **4,** 437–455 (2013).
75. Nayatani, Y. Simple estimation methods for the Helmholtz—Kohlrausch effect. *Color Res. Appl.* **22,** 385–401 (1997).
76. Brenner, E., Ruiz, J. S., Herráiz, E. M., Cornelissen, F. W. & Smeets, J. B. Chromatic induction and the layout of colours within a complex scene. *Vision Res.* **43,** 1413–1421 (2003).

77. Kingdom, F. A. A. Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Res.* **51,** 652–673 (2011).
78. Rudd, M. E. & Popa, D. Stevens's brightness law, contrast gain control, and edge integration in achromatic color perception: a unified model. *J. Opt. Soc. Am. A* **24,** 2766–2782 (2007).
79. Brill, M. H. & Carter, R. C. Does lightness obey a log or a power law? Or is that the right question? *Color Res. Appl.* **39,** 99–101 (2014).
80. Kuehni, R. G. Modeling lightness perception—another point of view. *Color Res. Appl.* **39,** 102–104 (2014).
81. Bouguer, P. *Traité d'optique sur la gradation de la lumière*. (1760).
82. Fechner, G. T. *Elemente der Psychophysik*. (Breitkopf und Härtel, 1860).
83. Ekman, Gös. Weber's Law and Related Functions. *J. Psychol.* **47,** 343–352 (1959).
84. Fairchild, M. D. *Color appearance models*. (Addison-Wesley, 1998).
85. Plateau, J. Ueber die Messung physischer Empfindungen und das Gesetz, welches die Stärke dieser Empfindungen mit der Stärke der erregenden Ursache verknüpft. *Ann. Phys.* **226,** 465–476 (1874).
86. Carter, R. C. & Brill, M. H. Modeling lightness perception—a response to Kuehni. *Color Res. Appl.* **39,** 105–106 (2014).
87. Stevens, S. S. Concerning the psychophysical power law. *Q. J. Exp. Psychol.* **16,** 383–385 (1964).
88. Marks, L. E. & Stevens, J. C. Individual brightness functions. *Percept. Psychophys.* **1,** 17–24 (1966).
89. Stevens, S. S. The Psychophysics of sensory function. *Am. Sci.* **48,** 226–253 (1960).
90. Stevens, S. S. The surprising simplicity of sensory metrics. *Am. Psychol.* **17,** 29–39 (1962).
91. Stevens, J. C. & Stevens, S. S. Brightness Function : Effects of Adaptation. *J. Opt. Soc. Am.* **53,** 375–385 (1963).
92. Jameson, D. & Hurvich, L. M. Theory of brightness and color contrast in human vision. *Vision Res.* **4,** 135–154 (1964).
93. Horeman, H. W. Relations between brightness and luminance under induction. *Vision Res.* **5,** 331–340 (1965).
94. Haubner, P., Bodmann, H.-W. & Marsden, A. M. A unified relationship between brightness and luminance. *Siemens Forsch. Entwicklungsberichte* **9,** 315–318 (1980).
95. Bodmann, H. W. Elements of photometry, brightness and visibility. *Light. Res. Technol.* **24,** 29–42 (1992).
96. Carter, R. Gray-scale perceptions calculated: optimum display background luminance. *Appl. Opt.* **36,** 1705–1717 (1997).
97. Bartleson, C. J. & Breneman, E. J. Brightness Perception in Complex Fields. *J. Opt. Soc. Am.* **57,** 953–956 (1967).
98. Fairchild, M. D. Considering the surround in device-independent color imaging. *Color Res. Appl.* **20,** 352–363 (1995).
99. Moroney, N. *et al.* The CIECAM02 color appearance model. in *Color Imaging Conf.* **2002,** 23–27 (Society for Imaging Science and Technology, 2002).
100. Luo, M. R., Gao, X. W. & Scrivener, S. a. R. Quantifying colour appearance. part V. simultaneous contrast. *Color Res. Appl.* **20,** 18–28 (1995).
101. Wu, R.-C. & Wardman, R. H. Proposed modification to the CIECAM02 colour appearance model to include the simultaneous contrast effects. *Color Res. Appl.* **32,** 121–129 (2007).
102. Marsden, A. M. Brightness—a review of current knowledge. *Light. Res. Technol.* **1,** 171–181 (1969).
103. ASTM E679 - 04(2011) Standard Practice for Determination of Odor and Taste Thresholds By a Forced-Choice Ascending Concentration Series Method of Limits.
104. Laven, P. *MiePlot*. at <www.philiplaven.com>
105. Lawless, H. T. & Heymann, H. *Sensory Evaluation of Food: Principles and Practices*. (Springer, 2010).
106. Fu, C. *et al.* An investigation of colour appearance for unrelated colours under photopic and mesopic vision. *Color Res. Appl.* **37,** 238–254 (2012).
107. International Standards Organisation. ISO 3591 Sensory analysis -- Apparatus -- Wine-tasting glass. (1977). at <http://www.iso.org/iso/catalogue_detail?csnumber=9002>

108. EPA. Method 180.1: Determination of Turbidity by Nephelometry. Rev. 2.0. *Methods Chem. Anal. Water Wastes* (1993).
109. Rossi, M., Gadia, D., Marini, D. & Rizzi, A. Towards image-based measurement of perceived lightness applied to paintings lighting. in *Proc. 6th Int. Symp. Virtual Real. Archaeol. Cult. Herit. VAST* 114–119 (2005). at <http://vpa2.sabanciuniv.edu/conferences/VAST2005/shortpapers/short2001.pdf>
110. Poynton, C. A. *Digital video and HDTV algorithms and interfaces*. (Morgan Kaufmann Publishers, 2003). at <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=349325>
111. Kinney, J. Comparison of Scotopic, Mesopic, and Photopic Spectral Sensitivity Curves. *J. Opt. Soc. Am.* **48,** 185–190 (1958).
112. Dadic, M. & Belleau, G. Beer Hazes. I. Isolation and Preliminary Analysis of Phenolic and Carbohydrate Components. *J. Am. Soc. Brew. Chem.* **38,** 154–158 (1980).
113. Hyams, D. G. *CurveExpert Professional v1.1.3*. at <www.curveexpert.net>
114. Marks, L. E. & Stevens, J. C. The form of the psychophysical function near threshold. *Percept. Psychophys.* **4,** 315–318 (1968).
115. O'Mahony, M. *Sensory evaluation of food: statistical methods and procedures*. (Marcel Dekker, 1986).
116. Moskowitz, H. R. Magnitude Estimation: Notes on What, How, When, and Why to Use It. *J. Food Qual.* **1,** 195–227 (1977).
117. Koo, T.-Y., Kim, K.-O. & O'mahony, M. Effects of Forgetting on Performance on Various Intensity Scaling Protocols: Magnitude Estimation and Labeled Magnitude Scale (green Scale). *J. Sens. Stud.* **17,** 177–192 (2002).
118. Kim, K.-O. & O'mahony, M. A New Approach to Category Scales of Intensity I: Traditional Versus Rank-Rating. *J. Sens. Stud.* **13,** 241–249 (1998).
119. Luo, M. R. & Hunt, R. W. G. Testing colour appearance models using corresponding-colour and magnitude-estimation data sets. *Color Res. Appl.* **23,** 147–153 (1998).
120. Moskowitz, H. R. Magnitude Estimation: Notes on What, How, When, and Why to Use It. *J. Food Qual.* **1,** 195–227 (1977).
121. Stevens, S. S. & Stevens, J. C. Brightness function: Parametric effects of adaptation and contrast in Program of the 1960 Annual Meeting of the Optical Society of America. *J. Opt. Soc. Am.* **50,** 1139 (1960).
122. Stevens, J. C. & Marks, L. E. Stevens power law in vision: exponents, intercepts, and thresholds. *Fechner Day* **99,** 82–87 (1999).
123. Li, X. & Gilchrist, A. L. Relative area and relative luminance combine to anchor surface lightness values. *Percept. Psychophys.* **61,** 771–785 (1999).
124. Park, J.-Y., Jeon, S.-Y., O'mahony, M. & Kim, K.-O. Induction of Scaling Errors. *J. Sens. Stud.* **19,** 261–271 (2004).
125. Debevec, P. E. & Malik, J. Recovering high dynamic range radiance maps from photographs. in *ACM SIGGRAPH 2008 Cl.* 31:1–31:10 (ACM, 2008). doi:10.1145/1401132.1401174
126. Szeliski, R. *Computer vision: algorithms and applications*. (Springer, 2011).
127. Sadar, M. J. Turbidity science. Technical Information Series—Booklet no. 11. *Hach Co. Loveland CO 7* **8,** (1998).
128. Moskowitz, H. R. Sensory evaluation by magnitude estimation. *Food Technol* **28,** 16 (1974).
129. Gilchrist, A. L. When does perceived lightness depend on perceived spatial arrangement? *Percept. Psychophys.* **28,** 527–538 (1980).
130. Krawczyk, G., Myszkowski, K. & Seidel, H.-P. Lightness Perception in Tone Reproduction for High Dynamic Range Images. *Comput. Graph. Forum* **24,** 635–645 (2005).
131. Gilchrist, A. *et al.* An anchoring theory of lightness perception. *Psychol. Rev.* **106,** 795–834 (1999).
132. Senol, S. Measurement system analysis using designed experiments with minimum α–β Risks and n. *Measurement* **36,** 131–141 (2004).
133. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Int. J. Nurs. Stud.* **47,** 931–936 (2010).

134. Hawkins, D. M. Diagnostics for conformity of paired quantitative measurements. *Stat. Med.* **21,** 1913–1935 (2002).
135. Camera HAL v3 overview | Android Developers. at <https://source.android.com/devices/camera/camera3.html>
136. CHDK Wiki. at <http://chdk.wikia.com/wiki/CHDK>
137. Camera raw, DNG : Digital Negative (DNG) | Adobe Photoshop CC. at <http://www.adobe.com/products/photoshop/extend.displayTab2.html>
138. The Programming Language Lua. at <http://www.lua.org/>
139. CHDK-Forum • Thema anzeigen - ptpCamGui - USB-Kamerasteuerung inkl. Datei-Download/Upload. at <http://forum.chdk-treff.de/viewtopic.php?f=7&t=2207>
140. Kim, S. J. & Pollefeys, M. Robust radiometric calibration and vignetting correction. *Pattern Anal. Mach. Intell. IEEE Trans. On* **30,** 562–576 (2008).
141. Decoding raw digital photos in Linux. at <http://www.cybercom.net/~dcoffin/dcraw/>
142. Rasband, W. S. *ImageJ, US National Institutes of Health, Bethesda, Maryland, USA*. (1997).
143. PHYS 2070 - Observational Astronomy - University of Manitoba - ImageJ Plugins. at <http://www.umanitoba.ca/faculties/science/astronomy/jwest/plugins.html>