2012

An Analysis of the factors influencing paper selection for books of reproduced fine art printed on digital presses

Brian M. Gamm

Franziska Frey

Susan Farnand

An Analysis of the Factors Influencing Paper Selection for Books of Reproduced Fine Art Printed on Digital Presses

By

Brian M. Gamm

Graduate Student, Chester F. Carlson Center for Imaging Science

Franziska Frey, Ph.D.

McGhee Distinguished Professor, School of Print Media

Susan Farnand

Staff Scientist, Chester F. Carlson Center for Imaging Science

Rochester Institute of Technology

A Research Monograph of the Printing Industry Center at RIT

No. PICRM-2012-02

R·I·T Printing Industry Center
An Alfred P. Sloan Foundation Center

An Analysis of the Factors Influencing Paper Selection for Books of Reproduced Fine Art Printed on Digital Presses

By

Brian M. Gamm
Graduate Student, Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

Franziska Frey, Ph.D.
McGhee Distinguished Professor, School of Print Media
Rochester Institute of Technology

Susan Farnand
Staff Scientist, Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

R·I·T Printing Industry Center
An Alfred P. Sloan Foundation Center

A Research Monograph of the
Printing Industry Center at RIT
Rochester, NY
November 2011

PICRM-2012-02

The research agenda of the Printing Industry Center at RIT and the publication of research findings are supported by the following organizations:

Table of Contents

Executive Summary 3
Introduction 4
Literature Review 6
Methodology 26
Results and Analysis 62
Conclusions 116
Addendum 119
References 120

Executive Summary

The purpose of this study was to understand and analyze the factors contributing to the selection of paper for books of reproduced fine art printed on digital presses. Toner-based digital presses are now capable of matching offset lithographic presses in both image and print quality. Current trends show that there is an increased interest in printing fine art books on digital presses. This research analyzed paper properties that maximize image quality and therefore influence the preference for digitally printed fine art reproductions. By extension, the papers rated as having the highest image quality and as being the most preferred would most likely be selected for use in books of reproduced fine art.

This research involved three stages:

1. Interviews with professionals involved in the production of fine art books.
2. A psychophysical experiment in which observers ranked images on the basis of image quality.
3. Physical measurements of the sample papers.

During the first stage, interviews were conducted with 13 professionals in fields including curatorial, publishing, printing, graphic design, and paper manufacturing. Questions included surveys of the roles different people play in the production process, the importance of different factors to paper selection, financing, and a general discussion of fine art reproduction.

During the second experimental stage, four images—representing four types of art media—were printed on twelve papers using two digital presses: an HP Indigo 7000 and a Kodak NexPress S3000. The twelve papers represented different combinations of coolness, print-show-through, roughness (Parker-Print Surf) and gloss (60 degrees). A psychophysical experiment was conducted in which observers ranked the prints on the twelve papers on the basis of image quality, color rendering quality, and surface appearance quality. The results were analyzed, and a model was developed to predict the probability that a paper was ranked in the top three. Model parameters included measurements of paper color (coolness), basis weight, roughness and gloss, gathered during the third stage. There was no previous metric for quantifying coolness, so additional experimentation was conducted to develop a model to predict the perception of coolness using colorimetry. An alternative experiment model was also developed that included parameters such as caliper, print gloss, line raggedness, and dot circularity. The resulting models allowed for the optimization of paper parameters that maximize the probability a paper will produce preferred, high-quality images.

One finding of this research was that the probability of a book being judged as having high image quality was optimized for papers with high coolness, low roughness and low gloss. Neither print-show-through, line raggedness, nor mottle were significant factors.

Another finding resulted from an additional lexical analysis that was performed for observer descriptions of their ranking behavior. This analysis provided complementary data to the psychophysical results. Observers' descriptions of their ranking strategies did not match the rank data, suggesting a possible disconnect between observers' conscious and subconscious ranking behaviors.

Introduction

"We want to be aware of the medium but not be ruled by it."
David Pankow, Director, RIT Cary Graphic Arts Press

The RIT Cary Graphic Arts Press, in Rochester, NY, published a book in 2009 entitled *Mother and Daughter, Posing as Ourselves*, by Elaine O'Neil and Julia Hess O'Neil (2009). This book is one of a new breed of fine art books printed entirely on a digital press. O'Neil and Hess further explored the bounds of digital printing by printing the book entirely in black and white, with three gray inks and a black ink in a process called GGGK. As stated by Sampat and Sigg, "The end result has, perhaps, even surpassed what might have been obtained from an offset lithographic press" (O'Neil, 2009).

Sampat and Sigg's statement echoes the basis of this work. Books of reproduced fine art are typically printed using offset lithography. Digital presses have only recently become able to produce images of equal quality to offset lithography. While digital presses are limited by the variety of substrate sizes and production speed, the advent of variable data printing and variety of paper grades available (Vogl, 2008) makes them increasingly valuable for short-run and print-on-demand workflows. In addition, individual presses offer special features, such as the spot color abilities of the HP Indigo and the dimensional ink of the Kodak NexPress, setting them apart from others.

The predominant issue that arises when comparing offset lithographic and digital presses is that of image and print quality. The earliest digital presses, primarily used for business graphics, lacked the resolution and image coverage to compete with offset lithography in the reproduction of images. As is often the case with new technologies, digital press manufacturers continuously improved their devices and are now able to contend with the print and image quality offered by offset lithography.

Digital presses have been embraced by businesses for the production of marketing and promotional materials, direct mail, transactional and business communications, and on-demand color books (Frey, Christensen, & DiSantis, 2006). The print-on-demand (POD) capabilities of digital presses opened the doors for companies such as Lulu® to provide print-on-demand services to consumers interested in low-cost self-publishing. However, while POD companies provide acceptable products for the average consumer, books of reproduced fine art require a higher level of care in printing and collaboration between the artist, publisher, printer, and curator (if applicable) than is offered by POD companies. While the average consumer is satisfied with a low-cost, efficient, and overall acceptable quality process, a higher level of quality standards for both the printing process and substrate are required for the reproduction of fine art.

Prior work has explored the differences between offset and digital papers (DiSantis, 2007; Evans & LeMaire, 2005; Frey et al., 2006; Vogl, 2008; White, 2007), image quality between offset and digital processes (Farnand, 2008, 2009; Farnand, Frey, & Anderson, 2009), and digital versus offset print quality in general (Chung & Rees, 2006; Farnand, 2008, 2009; Waite, 2003; Xu & Kellogg, 2007). However, to this author's knowledge little work has been published on the use of digital presses for the reproduction of fine art. This research explored factors relating to the choice of substrate and image quality properties of substrates used in digitally printed fine art books from the perspective of artists, printers, publishers, and users (see Figure 1). This research included three experimental stages: targeted interviews of people in each of the four subsets illustrated in Figure 1, psychophysical analysis, and physical measurement. The end result will be a first look into factors relating to the choice of papers for fine art book printing, image quality attributes of papers used in fine art book printing, and physical measurements of those papers used in the psychophysical experiment.

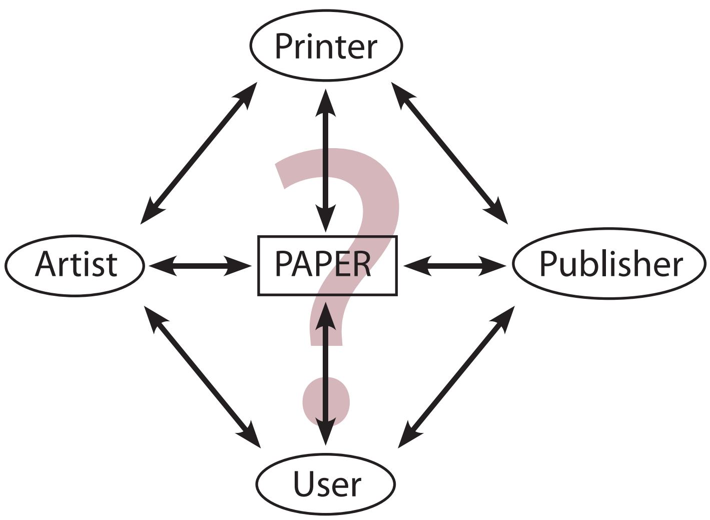


Figure 1. The hypothesized relationship between printers, artists, publishers, and users in the decision-making process for selecting papers used in books of reproduced fine art printed on digital presses

Literature Review

Electrophotography is not a new process. Chester Carlson invented the process in 1938, peddling his process with a demonstration kit in a cigar box. The process, rejected by corporations such as General Electric, IBM, and RCA, was first sold to the Battelle Memorial Institute in 1944, then later to the Haloid Corporation in 1947 (Notable Names Database, 2010). Haloid developed the first office copier, the 914, in 1959 and later changed its name to Xerox in 1961 to prepare for its initial public offering on the New York Stock Exchange (Xerox, 1999).

Chester Carlson's intentions were to simplify the process of copying documents. Since the release of the 914, the copying process has been identified by a simple button push. Secretaries were no longer needed to manually copy documents. In fact, Xerox's first commercial showed a man working in an office, handing his young daughter a document to copy. The young girl simply walks over to the copier, places the document on the scan bed and presses the copy button. Both Xerox and Canon (Japan) released color copiers in 1973 (Naudeau, 2002). The electrophotography industry quickly evolved with the advent of personal computers and the demand for high speed printing systems. However, as the consumer electrophotography industry flourished due to its efficiency and low cost, the quality of consumer systems was not within the bounds necessary for fine art reproduction.

There is not a large amount of published literature discussing papers used in digital fine art printing. However, since its inception in 2002, the Printing Industry Center at the Rochester Institute of Technology in Rochester, NY, has published research on differences between offset lithographic and digital press production (Chung & Rees, 2006; Farnand, 2008; Farnand et al., 2009), how substrates are used in digital printing applications (DiSantis, 2007; Evans & LeMaire, 2005; Frey et al., 2006; Vogl, 2008), and image quality comparisons between offset lithographic and digital presses (Farnand, 2008, 2009; Farnand et al., 2009). The following sections review these publications and other works discussing comparisons of the digital and offset lithographic print and image quality, digital press substrates, image quality analysis and standard substrate/print measurement.

Digital vs. Offset Quality

Printing technologies are often compared using both image quality metrics and print quality metrics. Image quality metrics are based upon human perception as determined by psychophysical experimentation. For example, a high-quality image would be one judged by human observers to have a higher quality than similar images. Engeldrum (1999, 2001, 2002, 2004a, 2004b) describes a hierarchical process for describing image quality in which general image quality—an integration of observer perceptions—is broken down into individual customer perceptions, such as hue, saturation, lightness, sharpness, contrast, roughness, etc. Print quality metrics were developed by the printing industry as quality control indicators. These are commonly measured using standardized methods and devices, and include physical measurements such as gloss, roughness, opacity, brightness, and color.

The division between offset lithography and digital press image and print quality is becoming ever more ambiguous as digital press technology becomes able to produce images of comparable quality to offset lithography. Chung and Rees (2006) investigated print quality differences between offset lithographic and digital technologies through a survey of over 150 printers operating both digital and offset presses. They found that:

> …the majority of color-related problems found within offset printing can be attributed to the materials involved in producing the printed product, whereas with digital print, color-related demerits appear to stem from the inherent constraints of the technology. (p. 30)

The offset lithographic mechanical process is highly repeatable. The same printing plates are used to print every page of a print run, the paper follows the same path for each print, and the press mechanics do not change through the course of a run. The problems encountered during an offset lithography run are largely due to control of the ink, fountain solution, substrate, and other consumable materials used throughout the run. On the other hand, the materials in a digital press are more controlled. Problems encountered during a digital press run are most often due to the press and the workflow (Chung & Rees, 2006).

A series of papers by Farnand (2008, 2009) investigating image quality differences between offset lithography and digital presses illustrates the ambiguity between image quality provided by the two technologies. During the transition from photographic film to digital sensors, the over-arching conversation among photographic consumers was when digital technology would surpass film technology in image quality. While digital photography has now overtaken film, print consumers have similar concerns. However, the current conclusions are not as cut-and-dry as in photography. Farnand had participants view images representative of several major categories of printing, including photo books, business graphics, advertising brochures, and test targets. The images were printed on uncoated and coated papers by one offset lithographic press and four digital presses. She found that offset lithography produced images with a greater than or equal perceived value on coated paper, while the opposite was true for digitally produced images on uncoated paper.[1] Participants pointed out that the offset printed images were more uniform and had higher quality lines, while the digitally printed images were of higher contrast (Farnand, 2008). Conversely, Farnand found a drop in the quality rating of digitally printed uncoated prints in her second study (Farnand, 2009). She also found that the variability due to media type was greater than the variability due to press type.

Although the jury is still out on whether the image quality of digitally printed documents can consistently equal that of offset lithography, studies have shown that

1 - However, the photo book and marketing materials printed on digital presses received higher ratings than those printed on the offset press.

the print quality produced by digital presses can equal, if not surpass, that produced by offset lithographic presses. Waite (2003) tested two digital presses and found that dot gain was minimal and consistent, the devices could handle large and small dot sizes, and that the solid ink densities were consistent across prints. However, both presses in question printed only black toner. Digital press technology has improved significantly in the seven years since that study was conducted.

Xu and Kellogg (2007) conducted print quality tests between a Heidelberg Speedmaster offset lithographic press and a Xerox iGen3 digital press. They analyzed CMYK solid ink density, print contrast, ink trapping, and dot gain. The reflection density was influenced more by the substrate than the press type. The iGen3 dot gain was lower than the Speedmaster, but less consistent. Overall print contrast and color gamut was higher for the iGen3. They concluded that the print quality of the iGen3 was comparable to that of the Speedmaster. On the other hand, a study by Rong (2009) compared print quality between the HP Indigo 3500 and offset lithography, and found offset lithography to have superior dot quality, line quality and resolution, and a wider color gamut.

Printing Paper Factors

Print Quality Requirements

Evans and LeMaire (2005) describe three main functional areas of paper performance: runnability, printability, and fitness for use. Runnability is described as "performance of papers in press operation, such that sheets will run smoothly through the print engine without jamming." Printability "relates to the image quality and overall appearance of the printed piece," and fitness for use is "the ability to be finished and distributed in the required manner, and the ability of the image to meet permanence requirements for the specific use" (p. 12). The overall performance of a paper during digital printing is largely determined by how well the paper performs during tone transfer and fusing.

Digital papers must be able to conduct some amount of electricity because the paper must attract the charged toner particles in order for them to adhere to the paper surface until fusing. If the paper is not able to evenly hold toner particles across its surface prior to fusing, problems such as mottling and poor image quality can occur. Uneven toner transfer is commonly the effect of factors such as spatial variations in the paper surface, paper thickness, moisture content, and filler concentration (Provatas Cassidy, & Inoue, 2004). Toner particles do not penetrate deeply into the paper surface. Electrophotographic systems rely largely on electrostatic force for tone transfer. Thus, those factors influencing tone transfer are affecting the electrostatic force between the paper and toner particles. Problems can occur on both a global and local level. Overall print density may be a result of global problems, while mottle, ghosting, and print density variations are commonly the result of local differences in electrostatic force (Kipphan, 2001).

Toner particles settle into small voids and pores in the paper surface during toner transfer. The toner particles adhere to the paper surface as a result of heat and pressure applied by the fusing rollers. Ideally, the toner particles evenly settle into the top paper surface during the tone transfer process, then fuse into that layer during the fusing process. The resulting print will have an even distribution of toner particles, and is relatively permanent and resistant to abrasion. If toner particles do not transfer well to the paper surface, do not fuse properly, or do not penetrate into the paper surface, then the print may be mottled and easily abraded (Evans & LeMaire, 2005). Papers with rough surfaces are notorious for poor toner transfer and fusing.

Permanence of Printed Documents

The factors affecting the permanence of a printed image depend upon the final use and storage conditions of the document. Mail-order documents are highly susceptible to abrasion, cracking, and light and air contamination in the automated sorting and mailing process (Frey et al., 2006). In addition, high humidity and water exposure are problems encountered in warmer environments. Of course, it is assumed that the printing environment itself is environmentally controlled. Otherwise, high humidity and temperature in the pressroom will cause errors in the printing process in addition to causing permanence issues.

Industry Use of Digital Press Papers

Evans and LeMaire (2005) surveyed 103 companies who provide digital printing services to investigate current paper grades used in digital printing and factors that determine paper brand and purchasing decisions. Marketing and promotional materials, quick printing applications, and direct mail comprised the top three most important jobs for digital printing. The most commonly used paper grades were coated gloss, premium uncoated, uncoated calendered, coated matte, uncoated uncalendered, and premium bond. When selecting a paper grade, print purchasers and print providers often collaborate on purchasing decisions. Runnability, print quality, availability of paper grades, appearance properties, price, multipurpose functionality and product range were important factors in the purchasing decision for a particular brand or source of paper. However, toner/ink adhesion, accurate sheet dimensions, dimensional stability, and moisture level had the greatest effect on the choice of a particular paper within a brand (Evans & LeMaire, 2005).

Moisture levels, accurate sheet dimensions and dimensional stability are important factors in maintaining the runnability of a paper. Uneven moisture levels across the paper surface will cause transfer issues and cause the paper to deform as water evaporates from the paper during fusing. Jams are often the result of poor dimensional consistency and out-of-plane deformations. However, due to the complicated feeding and fusing systems in digital presses, paper choices are limited to particular basis weights, sizes and thicknesses. A paper that is too thick will not be pliable enough to pass through the feeding system while a paper that is too thin may easily deform during the fusing process.

Vogl (2008) investigated the factors influencing the development of digital papers and digital presses. He found that dimensional stability and consistency of the product were critically important factors limiting performance and innovation of digital presses. However, current paper manufacturing technology limits the improvement of papers for digital presses. Toner adhesion was an important paper property limiting performance. Paper resistivity and moisture level were almost unanimously agreed upon as the most important factors influencing toner adhesion (Vogl, 2008).

There are trade-offs between paper characteristics throughout the printing process. For example, paper roughness facilitates paper feeding but inhibits tone transfer. High moisture levels help deter curling but reduce paper conductivity. While there are several important characteristics to consider in the production of a successful print, agreement on a specific set of characteristics is not universal across all areas of printing (Vogl, 2008). The choice of paper and important paper characteristics is largely application-dependent. Although the paper manufacturing process can be improved to solve most global problems, manufacturers themselves must produce paper tailored to the interests of several different groups within the printing community.

Quality Measurements

Organizations such as the Technical Association of the Pulp and Paper Industry (TAPPI), American Society for Testing and Materials (ASTM) and the International Standards Organization (ISO) publish standard test methods for commonly used paper and print measurements. ASTM and ISO publish standards for a wide range of industries, while TAPPI only publishes methods for the pulp and paper industry. Similar test methods are often mirrored between the three organizations. However, ASTM withdrew test methods for caliper, basis weight and stiffness because similar test methods were published by ISO and TAPPI. The following sections describe ASTM, TAPPI and ISO test methods for tests used in the physical measurement section of this thesis. Table 1 summarizes the latter test methods and lists their respective standards.

Table 1. Colorimetric and spectral error between measured and print predicted patches using ΔE_{00} and RMS optimization

Metric	TAPPI	ASTM	ISO
Surface Finish (Gloss)	T480 om-09 T653 om-98	D 523-08	2813:1994
Roughness/Smoothness	T555 om-09 T538 om-08	--	8791-4:2007
Caliper	T411 om-97	(D645M-97)	536:1995
Basis Weight/ Grammage	T410 om-98	(D645M-97)	536:1995
Whiteness	T562 pm-96	ASTM E313-05	--
Brightness	T452 om-98 T525 om-92	D985-97	2470-2:2008
Opacity	T425 om-01	(D589-97)	2471:2008
Solid Ink Density	--	D7305-08a	--
Print Gloss	D7163-05	--	19799:2007
Line and Solid Fill Quality	--	F1944-98	--

Paper Quality Metrics

Surface Finish (Gloss)

Surface finish, also known as gloss, is defined by TAPPI in two standards, T480 om-05 (TAPPI T480, 2003) and T653 om-98 (TAPPI T653, 2003). Specular gloss is defined in ASTM D 523-08 as the relative luminous reflectance factor of a specimen in the mirror direction. However, measurement angle of incidence varies depending on the standard and application. TAPPI T480 describes gloss measurement at 75 degrees illumination and detection. TAPPI T653 describes gloss measurement at 20 degrees illumination and detection. ASTM standard D523-08 and ISO 2813:1994 describe gloss measurements for 20 degree, 60 degree, and 85 degree measurements. Measurements of gloss from different angles can reveal important attributes of an object's surface. Objects with uniform surface characteristics should expect similar gloss measurements at different angles. However, specular reflectance (gloss) will vary at different angles due to irregularities of the object surface scattering light as the grazing-angle is approached.

Gloss measurement is often performed by gloss meters such as the BYK-Gardner micro-gloss and the BYK-Gardner micro-Tri-gloss.[2] The micro-gloss measures only at 60 degrees, while the micro-Tri-gloss measures at all three standard angles. Gloss meters are calibrated using a black glass tile with an index of refraction of 1.540 (TAPPI T480, 2003) for the sodium D line. This index of refraction differs slightly between standards. The specular reflection from the calibration tile is often given a gloss value of 100, relative to 100% specular reflectance and 0% scattering. The measurement systems are designed to correlate with the visual perception of surface shininess made at corresponding angles (ASTM D523, 2008).

Roughness/Smoothness

Roughness is defined in TAPPI T555 om-09 (TAPPI T555, 2003) as

> The mean gap between a sheet of paper or board and a flat circular land pressed against it under specified conditions. The mean gap is expressed as the cube root mean cube gap …

The roughness test measures the unevenness of a paper's surface by pushing air through the paper's surface. A rough paper will be less resistant to air flow through its surface than a smooth paper and will therefore result in a higher air pressure. Two TAPPI standards exist for the measurement of roughness: the Print Surf method (TAPPI T555, 2003) and the Sheffield Method (TAPPI T538, 2003). Each method uses a different measurement technique and records results in different units.

2 - Other brands are also available.

The Print Surf method device specifications are described in TAPPI T555 (TAPPI T555, 2003) as follows:

> The measuring head, which carries a circular metal measuring land surrounded by concentric guard lands, is pressed against the specimen, which is supported by a resilient backing surface consisting of lithographic blanket or other material designed to simulate packing materials used in printing processes. Air under pressure is led into the gap between one of the guard lands and the measuring land and the rate of flow between the edge of the measuring land and the specimen is measured.

An additional standardized method for measuring Print Surf roughness can be found in ISO 8791-4:2007 (ISO 8791-4, 2007). The Parker Print Surf, manufactured by Testing Machines Inc., is one common device for measuring print-surf roughness (see Figure 2a). The Print Surf measurement is recorded as the distance in micrometers between the paper sample and the flat circular land.

The Sheffield Method (TAPPI T538, 2003) uses an air leak tester to determine the smoothness of a paper surface. Air is pumped through a glass column at a constant pressure. The air is directed through a hole in the device's weighted head, which is then pressed upon the paper surface. The air flow through the glass column is directly affected by the resistance of the paper to the air passing through the device's head. A rough paper will allow more air to pass, thus increasing the air flow. The amount of air flow is measured by marking the position of a plastic float within the glass tube (see Figure 2d). Various tubes are used for different levels of air flow.

Figure 2. Devices are used to perform standard print quality tests:
(a) H.E. Messmer device for measuring Parker Print-Surf, (b) Technidyne Technibrite™ Micro TB-1C for the measurement of opacity, brightness, and various other color metrics, (c) Technidyne Brightmeter for the measurement of ISO Brightness, and (d) Device for measuring Sheffield Smoothness

Caliper

Caliper is the thickness of paper, paperboard or combined board. A standard method for measuring caliper is described in TAPPI T411 om-97 and ISO 536:1995. ASTM D645M-97 was withdrawn due to the existence of TAPPI and ISO methods. Caliper is measured using an automatically operated micrometer.

Basis Weight/Grammage

Basis weight, measured in pounds, is the United States standard measurement for paperweight. Grammage is a metric for describing paper density measured in g/m^2. Grammage is the common measurement for countries that use the metric system and is described in TAPPI T410 om-98 (TAPPI T410, 2003) and ISO 536 (ISO 536, 1995). The comparable ASTM standard, D645M-97 (ASTM D645M, 1997), was withdrawn due to the presence of both TAPPI and ISO standards. Basis weight is the weight of a 500-sheet ream with manufacturer specified dimensions. The dimensions of the ream vary depending on the type of paper. For example, bond paper has a standard ream dimension of 17 in. x 22 in. and offset paper has a standard ream dimension of 25 in. x 38 in. (Micro Format, Inc., 2010). An accurate measurement of grammage requires at least 5,000 cm^2 of paper surface area. Sheets are often cut to smaller sizes and stacked such that the total area is 5,000 cm^2.

Whiteness

Whiteness is defined in ASTM E313-05 (ASTM E313, 2005) as ”the attribute of color perception by which an object color is judged to approach the preferred white” (p. 971). Whiteness measurements are often described using two metrics: whiteness and tint. Whiteness metrics quantify the degree to which a white object appears white, and tint quantifies the redness or greenness of the white object. The most commonly used whiteness metric is CIE Whiteness, shown in Equation 1, often used in conjunction with CIE Tint, calculated using Equation 2,

$$\begin{gathered} W_{CIE} = Y + 800(x_n + x) + 1700(y_n - y) \\ 5Y - 280 > W_{CIE} > 40 \end{gathered} \tag{1}$$

$$\begin{gathered} T_{CIE} = 1000(x_n + x) - 650(y_n - y) \\ 3 > T_{CIE} > -3 \end{gathered} \tag{2}$$

where x_n and y_n are the chromaticity coordinates of the illuminant, x and y are the chromaticity coordinates of the sample, and Y is luminance.

TAPPI T562 pm-96 (TAPPI T562, 2003) and ASTM E313-05 (ASTM E313, 2005) describe methods for measuring and calculating CIE Whiteness and Tint. Paper whiteness and tint measurements are influenced by both the color and fluorescent properties of the paper. Paper additives, such as shading agents and optical brightening agents, in addition to the amount of bleaching and fillers in the paper pulp, can greatly affect the whiteness and tint of a paper.

Little research has been previously conducted on the perceptual relationship between paper whiteness and print quality. Coppel, Norberg, & Lindberg (2010) tested just-noticeable differences using pair-wise comparisons of papers with different shades and whiteness levels. They wanted to determine what whiteness difference was required for a just-noticeable-difference of image quality. They found that whiter substrates resulted in higher image quality ratings. Images were evaluated with a white paper border and without. Whiteness was compared to L^* and found to be more highly correlated to image quality than L^* when no paper surround was provided. However, the opposite was true; L^* was more highly correlated than whiteness to image quality when the paper surround was present.

Brightness

Brightness is an interesting standard because it is currently not used to measure the paper characteristic for which it was originally designed. Brightness was developed as a pulp bleaching process-control metric before the introduction of optical brightening agents into the paper industry (Axiphos GmbH, 2001). Brightness is standardized for both 45°/0° (TAPPI T452, 2003; ASTM D985, 1997) and d/0° measurement geometries (ISO 2470-2, 2008; TAPPI T525, 2003). The device (see Figure 2b) measures percent reflectance through a bandpass filter with a range from 395 nm to 515 nm and a peak at 457 nm ± 0.5 nm (ASTM D985, 1997). White objects appear brighter as they become bluer rather than yellower. When optical brightening agents became popular in the 1950s, brightness remained a popular metric. Over time, its purpose shifted from quantifying bleaching to quantifying fluorescence. While papers without optical brightening agents rarely have brightness values greater than 90, optical brightener fluorescence, which has a peak emission around 450 nm, can cause brightness measurements to surpass 95. Nevertheless, brightness is still a popular metric in the paper industry and is commonly used along with whiteness and colorimetric measurements to describe the appearance of white paper.

Opacity

Opacity is a measure of paper translucency. TAPPI T425 om-01 (TAPPI T425, 2003) defines the measurement of opacity with 15°/d geometry, illuminant A/2°, 89% reflectance backing and paper backing. Opacity is also defined in ISO 2471:2008 (ISO 2471, 2008). ASTM D589-97 (ASTM D589, 1997) previously defined a test method for opacity measurements, but was withdrawn due to available TAPPI and ISO standards. Opacity is calculated as

> 100 times the ratio of the diffuse reflectance, R_0, of a specimen backed by a black body of 0.5% reflectance or less to the diffuse reflectance 0.89, of the same specimen backed with a white body having an absolute reflectance of 89%. (TAPPI T425, 2003)

Opacity can also be calculated using a specimen "backed by a thick stack of the same kind of paper, R_∞" (TAPPI T425, 2003, p. 2). The method used to calculate opacity depends upon the application.

Opacity measurements are used to estimate the extent of show-through for a duplex print, an especially important attribute for book publication where every page is printed duplex. Show-through occurs when an image printed on one side of a sheet can be seen on the other side of the sheet when laid on a reflective white surface. This is especially important for books, where every page is printed duplex and many printed pages are stacked. Figure 2c shows the Technidyne Technibrite Micro TB-1C, a device commonly used for the measurement of opacity.

Print Quality Metrics

Solid Ink Density

Density is defined as the log of opacity, and is measured on solid ink patches printed with 100% of the individual process color. Process colors include cyan, magenta, yellow, and black.[3] Solid ink density (SID) is an important process control metric for presses. Changes in SID correlate to changes in image color. Changes in SID can result from many different factors, such as too much or too little ink (or toner), ink degradation, different half-toning algorithms, poor toner adhesion, and others.

Traditionally, SID is measured using a reflection densitometer. Different filter sets are used to measure visual, cyan, magenta, and yellow density. These filter sets are defined by a particular 'status.' Status T density is used in the United States, and is "applicable to the measurement of artwork for color separation and graphics arts materials such as ink-on-paper printed sheets, and off-press proofs" (ASTM D7305, 2008, p. 948). Status E density, used primarily in Europe, produces values more similar for all three chromatic inks at typical printing densities due to a narrower yellow filter. The latter filter set evolved from the wider of the two bandpass filter specifications in the German Institute for Standardization (DIN) document, DIN 16536-2:1986.

Spectrophotometers have begun to replace densitometers in recent years due to the burgeoning use of colorimetry in printing process control. In addition, density values can be calculated from spectrophotometer reflectance measurements.

Print Gloss

Print gloss measurement is standardized in ISO 19799:2007 (ISO/IEC 19799, 2007) and ASTM D7163-05 (ASTM D7163, 2005). Print gloss is measured on large areas of solid ink or tints using the same measurement method as for surface finish, described in Section 2.4.1.1. ASTM D7163-05 (ASTM D7163, 2005) designates 20 degrees, 60 degrees, and 85 degrees as measurement geometries for print gloss measurement.

Line and Solid Fill Quality

ASTM F1944-98 (ASTM F1944, 1998) describes techniques for measuring solid fill and line quality of printed images. Solid fill evaluations involve five analyses: mottling and coalescence, banding, bronzing, wet cockle, and dry cockle. Mottling is the non-uniformity of an image and density of a solid-fill area caused by the interaction between ink and paper. Banding is a defect consisting of alternating high and low density bands across a solid fill area. Bronzing occurs when the paper and black ink interact to produce an image with a bronze-like sheen. Cockle is unevenness in the paper surface caused by swelling of the paper fibers due to interaction with the ink. Wet cockle occurs when the ink is in the drying process, and dry cockle is what remains after

3 - If spot colors are used, they are also printed as solid inks.

the ink is dry. While solid fill defects may be difficult to detect in complex images, they may be especially noticeable in large areas of continuous tone.

The only line quality test defined in ASTM F1944-98 (ASTM F1944, 1998) is an image bleed line quality test. Image bleed occurs when inks bleed into regions reserved for other inks. For example, if cyan and yellow lines are printed adjacent to one another, the resulting image bleed will present itself as a green overprint between the two lines. This defect is usually due to the interaction between ink and paper. However, several tests exist to test the sharpness of a line. Image bleed can cause lines to overlap or appear blurred, and is influenced by dot size, ink viscosity and various paper properties. Toner is not as prone to image bleed as wet ink. However, toner particles may deform due to electrostatics during transfer and due to fusing pressure. In addition, lines constructed with toner particles may not appear as sharp as lines constructed with ink dots due to the large particle size. Software applications designed to measure line sharpness, aided by a scanning system, can be used to compute a metric for line sharpness (Quality Engineering Associates, Inc., 2009). Line quality measurements are especially useful for verifying that line quality is a significant determining factor in Customer Image Quality Rating, as discussed in research by Farnand (2008).

Image Quality

The Image Quality Circle

Engeldrum (2004b) defines image quality as the "integrated perception of the overall degree of excellence of an image" (p. 160). Image quality in the consumer-based realm (as opposed to strictly analytical, such as for military applications) is built upon a desire to understand why people prefer one image to another and what compel customers to purchase an image. Engeldrum developed a systematic approach for measuring image quality (Engeldrum, 1999, 2001, 2002, 2004b, 2004a). The Imaging Quality Circle, shown in Figure 3, outlines this approach.

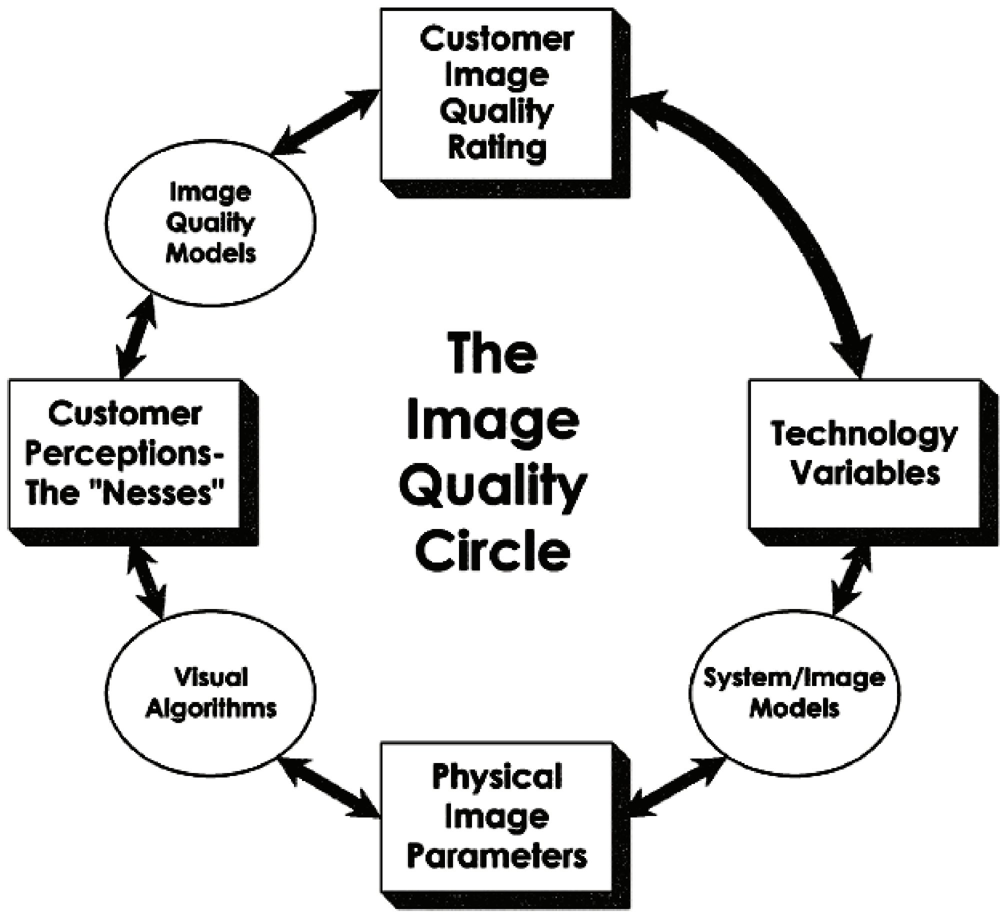


Figure 3. Engeldrum's Image Quality Circle (Engeldrum, 2004b)[4]

The Customer Image Quality Rating is the most general measurement of customer judgments; the customer rates the image using criteria they deem important. The investigator must be cautious to separate customer judgments and customer preference. Judgments are measures of image quality without bias to a particular aesthetic. Preferences are more subjective and describe what customers feel is aesthetically pleasing. For example, if a customer is asked to judge the best quality representation of a face, given the choices of a photograph, a Monet painting, and a Picasso, then they should choose the photograph as the highest quality representation due to the exactness with which the camera captures detail. However, the customer may prefer the Picasso based upon what the customer views as the most aesthetically pleasing representation. Perhaps the Picasso strikes a particular emotional chord. However, judgments are

4 - This approach provides the groundwork for conducting a psychophysical experiment analyzing image quality. The four components, are Technology Variables (i.e. paper manufacturing and print process parameters), Physical Image Parameters (i.e. paper and print quality metrics), Customer Perceptions (i.e. sharpness, colorfulness, contrast-ness), and Customer Image Quality Rating (general image quality). Visual Algorithms model the relationship between Physical Image Parameters and Customer Perceptions, and Image Quality Models model the relationship between Customer Perceptions and Customer Image Quality Rating.

not completely independent of preference, and it is important not to confuse the two. Judgments are commonly made prior to preferential decisions during experiments. It is necessary for the investigator to separate judgments and perceptions during the experiment, often using carefully crafted questions. Using the painting case as an example, a question leading into a judgment experiment may be, "Three images of faces are displayed before you. Which of these images has the highest quality representation of the scene? " A question leading into a preference experiment may be, "Which of the three images would you hang on your wall at home to look at every day?[5]" The relationship between the two criteria is commonly studied during analysis.

The other three major elements of the Image Quality Circle are Technology Variables, Physical Image Parameters, and Customer Perceptions. Technology Variables are the controllable elements of an imaging system, such as megapixels, dots per inch, substrate properties, etc. The number of Technology Variables in any given imaging system is immense. The simplest approach to studying image quality is to find a relationship between the changes in particular Technology Variables and Customer Image Quality Rating. However, Technology Variables are not what the customer sees when they view an image; instead, Technology Variables have a direct effect upon the Physical Image Parameters. Physical Image Parameters are those measurable aspects of an image, such as colorimetry, modulation transfer function, and optical density. Physical Image Parameters, like Technology Variables, are not directly related to Customer Image Quality Rating. However, they can be used to model Customer Perceptions, or, in Engeldrum's words, as the "nesses," since many sensations are suffixed with "ness." Customer Perceptions include colorfulness, sharpness, and graininess. Other terms, such as hue and chroma, are also "nesses," relating to the perceived hue and chroma of an image, but do not have the "ness" suffix. Customer Perceptions are the link between Customer Image Quality Rating and Physical Image Parameters, and it is important that models be derived to include this link.

There are three sub-features of the Image Quality Circle: System/Image Models, Visual Algorithms, and Image Quality Models, which are used to translate between the major elements of the Image Quality Circle. System/Image Models transform Technology Variables to Physical Image Parameters, and vice versa. For example, System/Image Models are used to relate the set of inks used in a printing system to measured tone reproduction, colorimetry, and contrast. Visual Algorithms are required to model the relationship between Physical Image Parameters and Customer Perceptions. The human visual system (HVS) does not respond linearly to changes in stimulus. The HVS is defined by its response to physical parameters such as contrast and color. Fairchild (2006) outlined color appearance models derived to relate physical measurements, such as luminance and spectral reflectance, to perceptions of brightness, lightness, hue, chroma, and other "nesses." He also described more complex models attempting to describe the entirety of human color response.

5 - It is assumed that a subject would only hang images in their home that they found to be aesthetically pleasing.

Image Quality Models represent the heart of the Image Quality Circle. The fundamental question in image quality is "Why do customers choose one image over another?" Customer Image Quality Rating is useful for identifying what images are preferred and by how much. Image Quality Models relate Customer Preferences to Customer Image Quality Rating, and help to identify why customers prefer one image to another (Engeldrum, 1999, 2001, 2002, 2004b, 2004a).

Thus, the most efficient and accurate process for relating Technology Variables to Customer Image Quality Rating is to follow the Image Quality Circle clockwise through Physical Image Parameters and Customer Perceptions, rather than attempting to directly relate Technology Variables to Customer Image Quality Rating.

Nevertheless, there is precedent for bypassing the measurement of Physical Image Parameters and Customer Perceptions. Farnand (2008, 2009) studied customer preferences of images reproduced using commercial offset lithography and digital presses on coated and uncoated papers. Farnand's research focused on relating Customer Image Quality Rating and the printing technology and paper type. While the amount of information gathered from this study was limited, the trends between response variance and preference could still be analyzed.

Every element of the Image Quality Circle used in an experiment requires experimental time and analysis. However, using all of the Image Quality Circle elements maximizes information gain and increases the chances of developing an accurate and physically meaningful model to predict Customer Image Quality Rating.

Image Judgment

Viewing an image is a highly complex task, as is vision in general. However, while the real world may be taken for granted as an accurate representation of itself, images of the real world are relatively abstract. Artistic renderings can vary greatly within the realm of abstractness, depending on the methods and style an artist uses to create the image. Most observers of artistic images, unless expert in the particular type of image, are left to interpret the images without insight into the artistic intent. The act of liking or disliking an image is an intuitive, yet complicated process that is taken for granted. However, the task becomes considerably more difficult when asked to explain why an image is liked, or, as in the case of image quality research, why one image is of higher quality than another. Vast amounts of research have been conducted in the fields of psychology, vision science, and marketing with the objective of understanding why people make decisions based on visual cues and how they communicate these decisions. This section will discuss research on the cognitive processes involved in viewing images and the lexicon observers use to make comparative decisions.

Visual Cognition

The act of viewing images is known to be dependent on both top-down and bottom-up cognitive processes. The image processing community weighs heavily on the bottom-up aspect of image quality because it is directly related to image saliency and is relatively easier to model than top-down processes such as attention and memory (Fredembach, Wang, & Woolfe, 2010). The raw rod and cone signal is first broken down into luminance and chrominance signals in the pathways leading to and within the lateral geniculate nucleus of the thalamus. After reaching area V1 of the occipital lobe, luminance information is broken down into orientation and spacial frequency contingent information. It is not until information is sent to the temporal and parietal lobes that object recognition takes place and judgments can be made (Wolfe, Kluender, Levi, Bartoshuk, Herz, Klatzky, & Lederman, 2006).

This process makes the bottom-up approach very attractive. Information is visually processed in components. Therefore, humans should be attentive to the most salient components of the image. Zeki (1999) emphasizes the functional segregation of the brain, especially in the visual system, and discusses in his book how art may be a subconscious result[6] of an artist's desire to isolate or emphasize specific visual elements. However, Zeki admitted that neurologists are still unsure of how the various processing elements of the brain integrate to provide us with a singular image of the world. What is known is that the visual systems provides humans with multiple consciousnesses: one of color, motion, face recognition, form, etc. Salient features are dominant because they emphasize a specific level of consciousness, and therefore it is possible that, when observers make image quality judgments, they are separately analyzing color, form, facial clarity, or other features confined to a single level of consciousness.

Fredembach and his coauthors (Fredembach et al., 2010) sought an answer to the question of whether observers judged image quality using a bottom-up or top-down approach. They tracked the eye movements of observers while the observers judged the quality of color images. Visual attention maps were derived from the average of observer responses. The visual attention maps were compared to several saliency prediction algorithms, including that by Itti and Koch (2000), what Fredembach and his coauthors consider the "de facto benchmark of saliency prediction" (p. 132). The researchers found that the saliency prediction algorithms performed worse than individual observer judgments when compared to the "ground-truth." While the saliency prediction algorithms are useful for random free-viewing situations, they were not as successful at determining where observers looked during image quality tasks. The authors also pointed out that artists often create images using tools such as local contrast, depth-of-field, and color balance to influence the attention of the observer. This further reduces the possibility that an observer randomly views a scene.

As observers view images during an image quality experiment, there are often additional cognitive influences other than salient factors. Memory plays an especially important role in image quality judgments when the stimuli are color images. Siple and Springer (1983) conducted a series of experiments to test color memory for fruits and vegetables under color-matching and color preference conditions. Observers were shown color photographs of the fruits and vegetables, then asked to reproduce the best

6 - Assuming the artist has little knowledge of neurology.

match and their most preferred match to the original on a visual colorimeter. Three different contextual cues were given, including object shape and object texture. They found that memory of hue was highly accurate, and that preference for Munsell hue[7] did not significantly differ from the original color. The same was shown to be true for Munsell value. However, the match and preference for Munsell chroma was significantly higher than the original object's chroma (Siple & Springer, 1983).

In many color-memory matching cases, human memory of colors are bounded by perceived naturalness. Naturalness is defined by de Ridder (1996, p. 489) as the "degree of correspondence between the reproduced image and reality." The utility of naturalness in memory relies upon the observer's past experiences and colors stored in long-term memory. Past work had shown a significant positive relationship between naturalness and image quality. This past work was expanded to include variations in chroma and lightness among the images. In de Ridder's experiments, subjects viewed four images at different chroma and lightness levels and rated the images on a scale from one to ten for perceived image quality, naturalness, and colorfulness (de Ridder, 1996). The researchers found that images rated as most colorful were also rated as the least natural. However, subjects commented that they were aware of this relationship between naturalness and colorfulness as they made their judgments. The researchers concluded that the relationship between naturalness and saturation and the relationship between image quality and saturation could be best described by an inverted u-shaped function.

Bodrogi and Tarczali (2001) use the idea of naturalness presented by de Ridder in their study of memory matching. The researchers deconstructed memory colors into two groups, "instant memory colors" and "later memory colors." The color stimulus at the point of detection by the visual system is referred to as the original color. After viewing a color stimulus, the perception of that color becomes an instant memory color. After a period of time, the memory color transitions from instant memory color to later memory color. The difference between these two memory colors was termed by the researchers as the 'memory shift.'

In their experimentation, Bodrogi and Tarczali had observers perform three tasks to test their memory of color. For the first task, observers adjusted a color patch displayed on a monitor to match their memory of a given object, such as grass. For the second task, observers viewed a color patch, waited for some time period, then constructed a match to that color on the display. For the third task, observers viewed a natural scene with a black rectangle isolating a solid patch of color from the scene, waited for some time period, then adjusted the color of isolated patch in a blurred rendition of the original image. Bodrogi and Tarczali had hypothesized that over time, memory of a familiar object would become confounded with "mean long-term memory color" (2001, p. 279. Those mean long-term memory colors seen often in the past were termed "prototypical

7 - Munsell hue is the categorization of color names within the Munsell Book of Colors, a color order system. According to Fairchild (2006), "The hue circle in the Munsell system is divided into five principle hues (purple, blue, green, yellow, and red…and is designed to divide the complete hue circle into equal perceptual intervals" (p. 97).

colors." They found that observers tended to shift their memory colors toward the mean long-term memory color by increasing its naturalness. The naturalness of a particular color is at a maximum when it is the prototypical color. If the original color is near to maximum naturalness, then the memory shift is minimized. A shift toward prototypical colors is aided by image context.

There is considerable research, as discussed above in this section, examining the prevalence of bottom-up and top-down processing in the completion of image quality experiments. Both types of processing play an important role in visual processing. Ramachandran and Blakeslee (1998) discussed very clearly the role bottom-up and top-down processing play in our visual system. In one example, they describe the act of seeing a cat:

> When we see a cat, its shape, color, texture and other visible attributes will impinge upon our retina and travel through the thalamus (a relay station in the middle of the brain) and up into the primary visual cortex for processing into two streams or pathways…one pathway goes into regions dealing with depth and motion…and the other to regions dealing with shape, color, and object recognition. Eventually, all the information is combined to tell us that this is a cat–say, Felix–and to enable us to recall everything we've ever learned about cats in general and Felix in particular.
>
> Now think of what's going on in your brain when you are imaging a cat. There's good evidence to suggest that we are actually running our visual machinery in reverse! Our memories of all cats and of this particular cat flow from top to bottom–from higher regions to the primary visual cortex–and the combined activities of all these areas lead to the perception of an imaginary cat by the mind's eye. (Ramachandran & Blakeslee, 1998, p. 109-110)

Yet, despite the constant activity of these two pathways, we do not perceive the object we are imaging as real. This is due to the constant sensory feedback from the retina to the higher levels of conscious processing (Ramachandran & Blakeslee, 1998). When we close our eyes, the retina produces a base-line signal informing the brain that no object is present. Likewise, when we have our eyes open, retinal signals overpower imagined objects. It is likely that this process is working with equal fervor when a observer is conducting an image quality experiment. The fact that an observer has the ability to conduct the experiment using anything more than random choice indicates that the observer has learned and committed to memory some aspect or aspects of image quality that they then apply while conducting the experiment. For experimental cases in which the observer either has possession of the original or has been shown the original images, it is likely that they are continuously recalling the memory of the original and making comparisons to the sample stimuli while conducting the experiment. In addition, the observer may also recall learned memories of image quality attributes. For experimental cases in which the observer was not shown the original, the observer is relying only on learned memory of image quality attributes and making comparisons to the sample stimuli. Of course, an observer's notions of the real world may have a drastic impact on their assessment of image quality, such as the case of naturalness judgments described by Bodrogi and Tarczali (2001).

Lexical Cognition

While the color processing path is similar in humans from the retina to the cortex, the way in which humans communicate color is unique for every individual. An individual's color lexicon is likened to a cognitive thesaurus containing the entirety of language that individual uses to describe color. Color naming is as much influenced by an individual's spoken language as it is by the socioeconomic environment in which they are raised (Beretta & Moroney, 2008). For example, more artistically-minded individuals are prone to describe their impressions of color rather than physical sensations, which might be the tendency for more scientifically-minded individuals. As such, artistically-minded individuals may be slower to name colors when provided such a task than scientifically-minded individuals (Beretta & Moroney, 2008).

Color Thesauri

The task of compiling a global color thesaurus has been undertaken on a number of occasions (Beretta & Moroney, 2008; Beretta, Moroney, & Recker, 2009; Moroney, 2003, 2009). There is broad industry interest in communicating digital color through natural language (Mojsilovic, Kovacevic, Hu, Safranek, & Ganapathy, 2000a; Mojsilovic, Kovacevic, Kall, Safranek, & Ganapathy, 2000b; Mojsilovic, 2002; Woolfe, 2007). Natural color language is especially useful in helping those naïve to color mathematics communicate color changes within a digital printing system. For example, if a person produced a print on their desktop ink-jet device and decided that the image needed to be 'warmer,' or 'a tinge more red,' they will input these natural language commands into a software program which will then translate the natural language into colorimetric changes and adjust the image accordingly. Experiments testing this technology are discussed by Woolfe (2007), and have been used as part of the Xerox Natural Langue Console in many printer drivers. Moroney created a website through HP Labs where observers could input the name of a set of randomly chosen colors (Moroney, 2003, 2009). Observers were allowed to participate as many times as they wished, and each time they viewed a different set of colors. Moroney developed a color thesaurus after analyzing the data from thousands of observers who participated in his web experiment. The Inter-Society Color Council and the National Bureau of Standards also collaborated on a dictionary of color names based upon the Munsell system, first published in 1955 (Kelly & Judd, 1976). However, this system has far fewer colors than Moroney's due to its strict organization within the Munsell system. Moroney's system allows users to input color names into a search bar and will provide synonyms and antonyms (based on colorimetric similarity and dissimilarity) for the given color name.

Mojsilovic (2002, p. 790) developed a color naming method using a perceptually based hierarchy of color vocabulary. Vocabulary were based upon the following syntax:

color name: achromatic name | chromatic name
achromatic name: lightness gray | black | white
chromatic name: lightness saturation hue | saturation lightness hue
lightness: blackish | very dark | dark | medium | light | very light | whitish

saturation: grayish | moderate | medium | strong | vivid
hue: generic hue | -ish form generic hue
generic hue: red | orange | brown | yellow | green | blue | purple | pink | beige | olive
-ish form: reddish | brownish | yellowish | greenish | bluish | purplish | pinkish

where : denotes "is defined as" and | denotes "or." A mathematical model was developed based upon this system.

Natural Language of Perceptual Attributes

Mojsilovic published a series of articles in which she studied the use of multidimensional analysis to determine the vocabulary and grammar used to describe color patterns (Mojsilovic et al., 2000a, 2000b; Mojsilovic & Rogowitz, 2001; Mojsilovic, 2002). She discusses how humans judge similarity only within the confines of a particular area of interest as opposed to globally and at random (Mojsilovic et al., 2000a). The goal of her research was to "detect basis visual categories that people use in judgment of similarity, and then design a computational model which accepts one (or more) texture images as input" (Mojsilovic et al., 2000a, p. 39). In her experiments, subjects viewed pairs of images of patterns with varying color and texture and rated the similarity on a scale between 0 and 100. Half of the participants were asked to describe why they made their decisions as such. The data was analyzed using weighted multidimensional scaling (Kruskal & Wish, 1976). Four prominent dimensions were determined: overall color, color purity, regularity and placement, and directionality. Overall color was defined as the "presence/absence of a dominant color" (Mojsilovic et al., 2000a, p. 41). A dominant color can either be a single color or a color seen as more dominant due to differences in saturation, contrast, or intensity. Purity was defined as the degree of colorfulness or as the chroma and saturation dimension. Mojsilovic created several grammar rules based upon cluster analysis of the image similarity. The separation between chrominance and luminance dimensions is due to the separation of chrominance and luminance channels in the human visual system. Most texture and frequency information is processed using the luminance channel of the visual system (Mojsilovic et al., 2000a, 2000b). Rao and Lohse (1996) identified natural language dimensions specific to texture discrimination in earlier work. They identified factors including contrast, repetitiveness, granularity, randomness, smoothness, density, and directionality as important factors in texture discrimination.

Montag and Kasahara (2001) applied multidimensional scaling methods to understand colorimetric factors important to image quality. In their experiment, observers viewed paired comparisons of seven prints. The prints were made on the same paper for five printers. Three different papers were used for one of the printers. Observers judged both preference and image quality for each pair, and plotted various color metrics against the primary dimensions. The multidimensional analysis revealed the apparent background color dimension and skin color dimension as important perceptual attributes in the image quality analysis.

Methodology

Interview Stage

Thirteen interviews were conducted with members of the printing, paper manufacturing, museum publication, curatorial, and graphic design industries. The purpose of these interviews was to understand current trends in paper selection for books of reproduced fine art and the associated paper property considerations. A series of prepared interview questions was asked to each interviewee. The questions were aimed at obtaining both specific and open responses. Specific responses revealed papers used, important paper properties, paper selection decision-making criteria, and any other information relevant to the paper selection process. Open responses allowed the interviewee to elaborate on areas such as differences in paper selection for reproductions of different art media, economic considerations, and relationships between printers, publishers, artists, and users. The interviewees and interview questions are discussed in the following sections.

Interviewees

The interviewees represented a wide range of fields involved in the process of publishing books of reproduced fine art. Three interviewees were curators; six worked in museum publications—small, internal documents such as brochures, annual reports, marketing, etc.—and/or publishing larger documents such as books; one worked in paper manufacturing; two worked in printing; and one worked in graphic design. The interviewees' employers included six museums, one graphic design firm, one library, one paper manufacturer, and two printing companies. Eight interviewees were located in Rochester, NY and interviewed in person. The remaining five interviewees did not work in Rochester and were interviewed via Skype. All interviews were recorded with consent from the interviewees. The five Skype interviews were recorded using Ecamm Call Recorder v2.3.10. The remaining in-person interviews were recorded using Audacity 1.3.12-beta and an Apple MacBook Pro laptop. Each interview was transcribed and analyzed for this research. All interviewees are identified by anonymous codes in this study. Table 2 shows the coded names of each interviewee along with their employer and job description.

Table 2. Coded names of each interviewee along with their employer and job description

Name	Employer	Job Description
Cur 1	Rochester Area Museum	Curator
Cur 2	Rochester Area University	Professor
Cur 3	Rochester Area Museum	Curator
Pub 1	New York City Area Museum	Director of Print On Demand/ Photographer
Pub 2	New York City Area Museum	Production Director
Pub 3	Los Angeles Area Museum	Production Manager/ Production Coordinator
Pub 4	Rochester Area Museum	Director of Publications
Pub 5	Smithsonian Institution	Deputy Publisher and Publications Director
Pub 6	Rochester Area University	Press Director and Curator
Pr 1	Rochester Area Printer	VP of Sales and Marketing
Pr 2	Rochester Area Marketing Firm	Director, New Business Development
GD 1	Rochester Area Design Firm	Graphic Designer
PM 1	Mid-Sized Paper Manufacturer	Director, Research and Development /National Business Director

Interview Questions

The interview questions were divided into three sections: demographic questions, direct research questions, and open-ended research questions. One of the original goals for the interview process was to understand the differences between people involved in different aspects of production in their contributions to paper selection. Four general categories were initially identified: publishing, printing, artists, and users. However, these categories were adjusted as the interview process evolved. The original interview questions are presented below in Table 3.

Table 3. Targeted interview questions

Type	No.	Question	Answer Choices (If applicable)
Demographic Questions	1	Would you classify your company or field as publishing, printing, an artist, or a user of reproduced fine art?	
	2	Please describe your position in your company or (if an artist) the medium in which you most frequently work.	
	3	How many years has your firm (or yourself) been active?	
	4	How many employees are currently in your company?	
	5	Which of the following best describes your company's 2004 revenues?	< $3 million $3 million to $5 million > $5 million to $10 million > $10 million to $15 million > $15 million to $20 million > $20 million Don't know Refused to answer

Type	No.	Question	Answer Choices (If applicable)
Direct Research Questions	6	Which digital presses do you use most often for books of reproduced fine art? If you do not use digital presses, which offset presses do you use?	
	7	Which properties of unprinted paper are important to your decision to select a particular paper for use in books of reproduced fine art? Please rate each property on a scale from 1 to 4, where 4 is "critically important" and 1 is "not important."	Grain Direction
			Paper Gloss
			Caliper
			Basis Weight
			Whiteness
			ISO Brightness
			Opacity
			Stiffness
			Other
	8	Which properties of printed paper are important to your decision to select a particular paper for use in books of reproduced fine art? Please rate each property on a scale from 1 to 4, where 4 is "critically important" and 1 is "not important."	Neutral Color
			Print Contrast
			Tone Reproduction
			Solid Ink Densities
			Print Gloss
			Differential Gloss
			Line Quality
			Other
	9	What are the paper brands (manufacturers) you use most frequently?	
	10	From the brand you use the most, which paper grades do you most frequently purchase for non-fine-art reproduction jobs?	
	11	From the brand you use the most, which paper grades do you most frequently purchase for fine-art reproduction jobs?	
	12	Of those papers you listed in the previous question, which do you use most frequently for fine art reproduction?	
	13	Who commonly decides which paper to select for printing the book of reproduced fine art? Please indicate the influence that each group of people has in the purchasing decision on a scale from 1 to 4, where 4 is "all of the weight" and 1 is "none of the weight."	Printer
			Publisher/Curator
			Artist (if living)
			User/ Expert Consultant (if artist is not living)
	14	What role do the following people bring to the process of deciding which paper to use in a book of reproduced fine art? Please indicate the influence that each group of people has in the purchasing decision on a scale from 1 to 4, where 4 is "all of the weight" and 1 is "none of the weight."	Printer
			Publisher/Curator
			Artist (if living)
			User/ Expert Consultant (if artist is not living)
Open Ended	15	Please explain what influence the original artwork's medium has on the choice of paper for reproduction.	
	16	Please explain how the integrity of the reproduction is influenced by the choice of paper.	

Psychophysical Experiment

It is impossible to achieve a quantitative study of human judgment without a designed psychophysical experiment. This study was rooted in the principles of the Image Quality circle [Engeldrum04a] which provides a method for modeling human response based upon physical measurement. Observers from the experiments in this study first ranked images on the basis of image quality, then ranked them on the basis of color rendering quality and surface appearance quality. The latter two rankings were akin to Engeldrum's Customer Perceptions, or "nesses," while image quality is akin to Customer Image Quality Rating. The sections below describe the processes by which samples were selected, prepared, and printed, followed by the psychophysical design and the practical methods for running the experiment.

Method

Sample Selection

There are many different metrics by which papers can be categorized, including metrics used by paper manufacturers and metrics used by printers. The original plan was to select papers using a 2^k full factorial design, where k is the number of factors used to qualify the papers, and 2 is the number of levels within each factor by which a paper could be classified. Paper manufacturers rarely report quality control metrics, such as Parker Print Surf (ISO 8791-4, 2007) for roughness measurements and 60 degree gloss for gloss measurements, in their promotional materials. In many cases, customers do not care about these values or have their own means of making these measurements. Most information released by paper manufacturers to consumers about their papers includes brightness, opacity, basis weight, and thickness, in addition to information about post-consumer-waste content and other environmental considerations. Furthermore, not all paper companies provide this complete list of measurements, and, when they do, it is not necessarily accurate.

Therefore, the first obstacle encountered in selecting papers for this study was that they could not be selected using quantitative measurements. The paper market contains may different paper brands. It would be a daunting task to collect all of these samples and measure them to determine the proper paper samples for measurements. Therefore, a practical concession was made for sample selection. Four factors were identified by which paper could be selected using visual and tactual methods. Those factors were roughness, gloss, opacity, and color, all of which are easily distinguishable either by looking at or feeling the potential paper samples. Each factor was categorized into the following levels to satisfy the 2^k full factorial sampling structure: rough and smooth for roughness, high and low for gloss, high and low for opacity, and warm and cool for color.

Theoretically, 16 papers where needed to fulfill the 2^k full factorial sampling structure (two levels per factor with four factors requires 16 papers). However, it was determined, based upon conversations with representatives of paper manufactures and paper distributors, that papers containing both high roughness and high gloss were not manufactured. Therefore, the four elements of the full factorial design containing rough and glossy papers were removed. Table 4 shows the 12 combinations of color, PST, roughness, and gloss that were included in the study.

Table 4. The 12 papers used in the psychophysical experiment[8]

Paper Code	Color	Opacity	Roughness	Gloss
A	cool	low	smooth	low
B	cool	high	smooth	low
C	cool	low	rough	low
D	cool	high	rough	low
E	cool	low	smooth	high
F	cool	high	smooth	high
G	warm	low	smooth	low
H	warm	high	smooth	low
I	warm	low	rough	low
J	warm	high	rough	low
K	warm	low	smooth	high
L	warm	high	smooth	high

The sample selection was originally designed to keep basis weight constant. However, it was determined during sample selection process that there was not enough noticeable variation in opacity among papers of the same basis weight. Therefore, basis weight was included as a confounding factor with opacity. All papers with low opacity were 80lb text weight, and all high opacity papers were 100lb text weight.

Several factors were considered during the process of selecting papers to fill the criteria in Table 4. First, during the Interview Stage, interviewees frequently listed papers they commonly used in production. While it was not necessary to use only papers listed by interviewees, several papers were included that had been used previously in commercial production. Papers E and F were the same brand of paper differing by basis weight. That same brand had been used in a book published by the Memorial Art Gallery (MAG), in Rochester, NY, entitled *Seeing America*, documenting the American art collection held by the MAG (Searl, 2006). Papers A, B, and J were from the same family of papers used recently in an RIT Cary Graphic Arts Press book entitled *Mother and Daughter Posing as Ourselves*, by Elaine O'Neil and Julia Hess (2009). The second consideration was simply to find papers that fit each of the 12 conditions. This task proved considerably difficult. The staff of the Rochester, NY XpedX sample room provided much help. They provided sample books of several different brands and made recommendations of papers that would fit the experiment's specifications. The final consideration was the acquisition of papers. Nine of the twelve papers were manufactured by Mohawk Fine Papers. Mohawk Fine Papers donated paper to the study for each of the nine Mohawk papers selected for the study. The remaining three papers were purchased from a distributor. Table 5

8 - Each contained a unique combinations of color, PST, roughness, and gloss. High and low values for each parameter were selected as the basis on which papers were selected, similar to a factorial design.

shows the 12 papers used in this study along with the experiment code associated with each paper. Each of the 12 papers represents the set of paper properties listed in Table 4, respective to the paper code.

Table 5. Twelve papers selected to represent the selection parameters listed in Table 4 by experiment code

Paper Code	Papers
A	Mohawk Superfine Smooth Ultrawhite 80lb Text iTone
B	Mohawk Superfine Smooth Ultrawhite 100lb Text iTone
C	Mohawk Via Felt Bright White 80lb Text iTone
D	Neenah Coronado Infinite White Stipple 100lb Text
E	Burgo Chorus Art Digital Gloss 80lb Text
F	Burgo Chorus Art Digital Gloss 100lb Text
G	Mohawk Beckett Expressions Candlelight 80lb Text iTone
H	Mohawk Beckett Expressions Candlelight 100lb Text iTone
I	Mohawk Via Felt Warm White 80lb Text iTone
J	Mohawk Superfine Eggshell 100lb Text iTone
K	Mohawk Loop Silk Coated 80lb Text iTone
L	Mohawk Loop Silk Coated 100lb Text iTone

Test Targets and Sample Design

This experiment required two digital press runs to produce test targets and experiment samples. The purpose of the first press run was to produce targets from which print quality attributes could be measured and for use as profiling targets for the second press run. Several test targets were printed during the first press run to provide a variety of different measurements. The purpose of the physical print quality measurements was to provide potential Physical Image Parameters for the psychophysical models. While many different print quality attributes could be used as predictors in this model, it was important to choose only those attributes that could logically predict the results. In addition, several predictors may be highly correlated. In this case, only one of the highly correlated predictors would be used, as much time and effort would be required to measure all possible predictors. Several predictors were chosen that the experimenters felt would illuminate important aspects of print quality and have minimal confounding amongst themselves. Those factors were: 0.1 mm dot circularity, line raggedness, 40% print mottle, print gloss (100% CMY), and gloss differential. Two targets were printed from which these measurements were made. The targets shown in Figures 4 and 5 were printed without embedded profiles.

Figure 4 was acquired from the web site of Quality Engineering and Associates (QEA). The target was designed for use with QEA's Image Analysis System Lab (IASLab™) software (Quality Engineering Associates, Inc., 2009). IASLab™ Version 2.12.4.0 was used in this thesis. According to the materials provided by QEA, "IASLab™ is an advanced software product for objective, automated evaluation of image quality." The IAS Test Target was designed for use on any of QEA's IAS products.

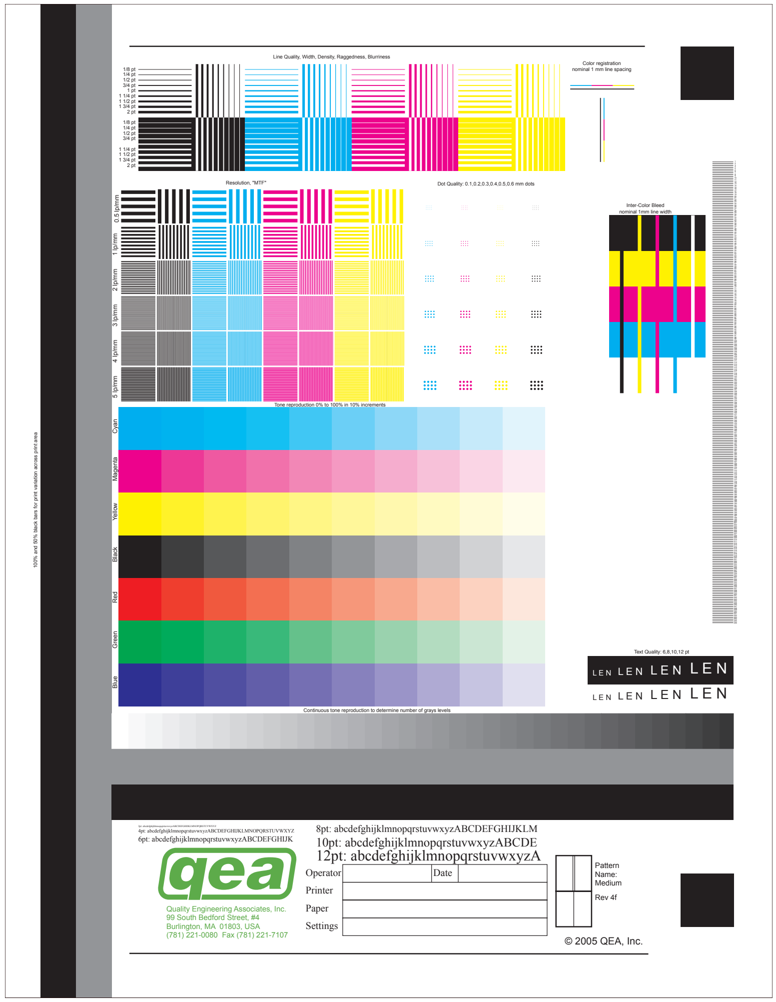


Figure 4. QEA test target

Figure 5 was originally designed for mottle measurements. However, the large surface area of the patches made this target ideal of the measurement of print gloss. Furthermore, it was found that the solid patches in the IAS Test Target could also be used for mottle measurements.[9]

9 - While mottle would have best been measured using a large-surface area target, the ability to integrate the mottle measurements into an automated workflow with line raggedness and dot circularity using IASLab™ was more beneficial to the process than using a separate target.

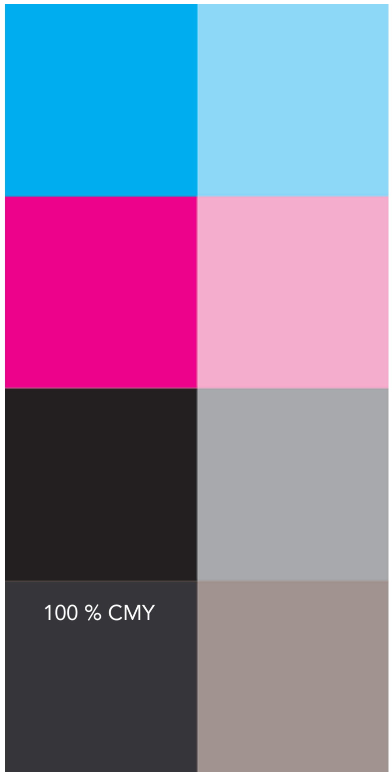


Figure 5. Test target containing solid area and 40% cyan, magenta, and black patches

Four copies of the IT8.7/4 test target (shown in Figure 6) were printed for each of the 12 papers to be used in generating profiles. This target is one of a group of targets standardized by the American National Standards Institute (ANSI). IT8.7/4 targets can be formatted for measurement by many different devices. All targets printed during this experiment were measured using the X-Rite iSis XL scanning spectrophotometer. The diamond markers on the sides and the black bar at the top of the targets are positioning markers used by the iSis for target registration. The black and white squares directly boarding the color test patches are measurement markers used by the now discontinued X-Rite DTP70 Autoscan Spectrophotometer (X-Rite, 2011), a predecessor to the iSis. These regions of the image were not used in this experiment.

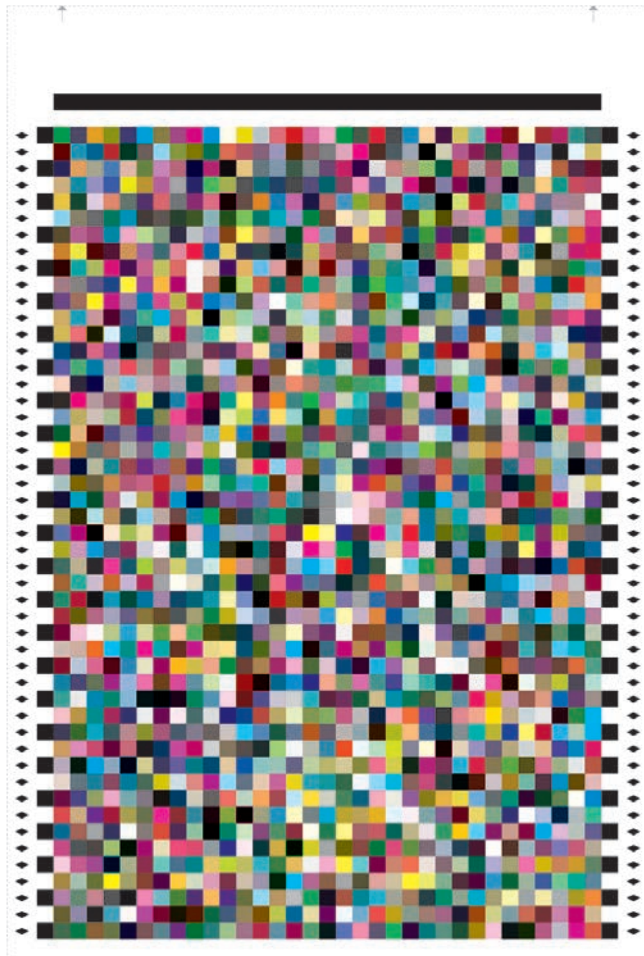


Figure 6. IT8.7/4 Target formatted for the X-Rite iSis

The experiment samples were printed during the second press run. Figure 7 shows the four images used in the experiment. Each image represents a different art medium. Figure 7a is an aquatint print, Figure 7b is an oil painting, Figure 7c is a sepia platinotype photograph from the archives of the RIT Image Permanence Institute, and Figure 7d is a watercolor painting.

a. b.

c. d.

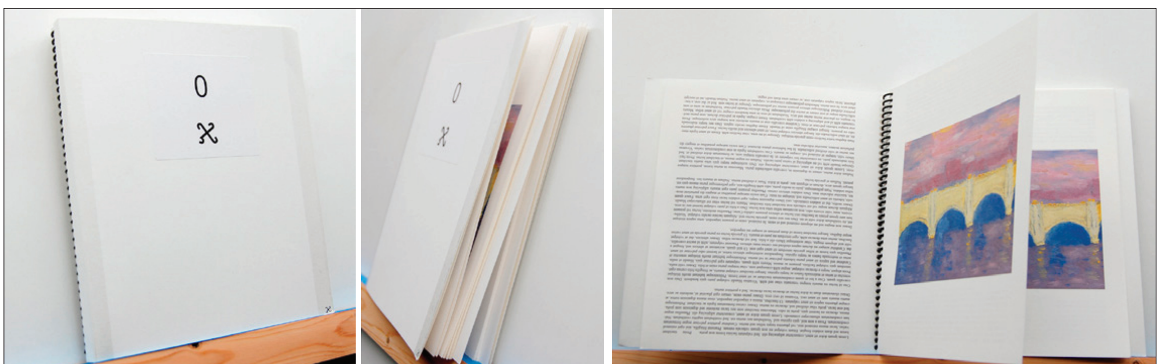
Figure 7. Images contained in the experiment book: (a) aquatint print, (b) oil painting, (c) sepia platinotype photograph, and (d) watercolor painting

The particular photographic reproductions were chosen from among the reproductions used in the *Benchmarking Art Interchange Cycles* project, funded by the Andrew W. Mellon Foundation (Frey, Farnand, & Jiang, 2010). In this research, participating museums were provided the original artwork and asked to create digital photographic reproductions. The files were printed using a common offset lithography workflow.

Frey, Farnand, and Jiang conducted two sets of psychophysical experiments to determine which workflow produced the highest image quality. The original artwork was present during one experiment set and absent from the other. The particular digital reproductions used in this experiment were rated highly in both experiment sets. The selection of highly rated images minimized the chance that the digital reproduction's quality might affect observer judgments during the psychophysical experiment.

A sample book was created (Figure 8) using four repeated prints of the images shown in Figure 7. Each image was centered on the page after including a 0.5-inch gutter on the left. A block of Lorum Ipsum text, divided into five return-separated paragraphs, backed up each image on the reverse side of the sheet.

The text was centered on a space within 0.75-inch margins on the left and right, 0.5-inch margins on the top and bottom, and an additional 0.5-inch gutter on the left. The book had dimensions of 8.5 inches by 10 inches, and was organized into four sections, one for each image. Each section contained four replicates of its respective image followed by ten unprinted sheets. The four image replicates served to prevent any visual changes of the image that might otherwise be caused by a different image on the next page, and provided replacements for any images that were damaged during the experiment. The ten unprinted sheets added bulk to the sample books so that they were similar in size to a commercially printed book. Between each section was a sheet of 100lb cover-weight paper. The cover was designed as a wrap-around and used the same cover-weight paper as was used to divide the sections. The book was collated and bound using a 0.5-inch black spiral bind. The binding was performed at the RIT Print and Postal Hub. A label was placed on the cover of each book containing the press code (0 for HP Indigo 7000 and 1 for the Kodak NexPress S3000) and the paper code. The papers were originally coded using the English alphabet. However, the paper codes were written on the books using the Georgian alphabet so observers could not recognize the coding system. Figure 8 shows three views of the sample books.



a. b. c.

Figure 8. Views of the sample books: (a) front cover with label containing the codes, (b) side view, and (c) open view with back-up text and image imposition shown

One book was printed for each press and paper combination, resulting in 24 books. The books were stored in a file cabinet in the experiment room to avoid excess light exposure when not in use.

Printing Process and Color Management

The complexity of this study is well illustrated by the large variety of workflows that could be used to print the experimental samples. As it is, there are 12 different papers, differing from all other papers by at least one of four factors: color, basis weight, roughness, and gloss. The goal of this thesis was to determine those paper factors that optimized image quality of fine art reproduction on digital presses. This entailed running a study where the image quality only differed due to the substrate and was independent of the printing process. There were several possible methods for achieving consistent printing across all sheets. The two most plausible methods differed in philosophy and outcome.

Image quality experiments can be designed to emulate the real world or exist purely as laboratory tests. Theoretical research experiments are often building blocks for natural case studies. Experiments conducted to emulate the real world prepare samples as they would be prepared commercially and depict subjects as seen in commercial products. The stimuli produced for experiments emulating the real world must be created using methods common to industry, and the images chosen as stimuli must represent commonly printed scenes. Without one of these components, the test can only be classified as theoretical research. For example, it is possible to study image quality by having observers judge large uniform color fields printed using a commercially viable process. However, people rarely look at uniform fields in commercial work. Likewise, stimuli could contain commercially viable images but be produced using nonstandard methods, only viable in the laboratory.

While these discussions of experimental procedure may seem elementary, they did, in fact, arise out of real discussions the experimenters had with the printing and pre-press departments at the RIT Printing Applications Laboratory (PAL). PAL was employed to print the samples for this study on the Kodak NexPress S3000 and the HP Indigo 7000. However, the printing process was complicated. Images had to be printed on the 12 different paper stocks with minimal variability across substrate due to printing such that observers would only judge the images based upon differences in paper, and not be biased by differences in print quality. The discussion then turned to the definition of consistency. What did it mean to have consistency throughout the printing process across papers? It is here that the differences between design as theoretical research and design as a real-world study were differentiated. In a strict printing sense, understanding the absolute difference between papers without regard for the printing process would require printing the same CMYK across 12 papers. This would result in consistent ink distribution but potentially drastic differences in image appearance and print quality. One of the main questions relative to this approach was how to set the total area coverage. For example, PAL calibrated the HP Indigo 7000 using NewPage Sterling Ultra Digital paper. This is a coated paper and can achieve a maximum density much higher than an uncoated sheet. Laying the same amount of ink on each paper does achieve the goal of consistency across sheets, but may in fact introduce other consequences. Printing an uncoated sheet as if it were a coated sheet would result in differences in maximum density and inconsistencies in tone reproduction. This, in turn, could result in such large differences in print quality that the observer would judge the images based upon the merits of both the printing process and the paper rather than the merits of the paper alone.

The alternative to printing consistent CMYK across sheets is to use an ICC-based workflow with a relative colorimetric rendering intent to achieve colorimetric consistency but preserve the integrity of paper white. The result would be different CMYK mixtures for each paper, and printed images optimized for their respective papers. Image appearance would be consistent with respect to paper.

The difference in the printing process would also cause a difference in the objective of the psychophysical experiments. If consistent CMYK was the printing method, then the observers would be judging which paper resulted in the best image quality as a result of the total area coverage and the press. This is far from the process used in commercial digital printing, which is heavily dependent upon ICC color management. By using ICC color management to achieve consistent colorimetry, observers would be judging which paper resulted in the best image quality using a workflow to achieve the best possible image quality on each paper. This is more akin to a real-world workflow.

Here is one possible scenario that could be delivered to an observer:

> You just finished walking through an art gallery and now visit the gallery gift shop with the intention of purchasing a book of reproduced fine art. The Museum published the 12 books in front of you on 12 different paper stocks. Please rank the books in order of image quality.

This now seems like a viable scenario. There would never be a case where a publisher uses the same CMYK combinations for 12 papers so different in physical properties. However, it is also unlikely that a single printer would print the same book on 12 different substrates and take the time to color-manage each substrate. Therefore, the following scenario was selected for this experiment:

> You just finished walking through an art gallery and now visit the gallery gift shop with the intention of purchasing a book of reproduced fine art. Twelve publishers produced the same book. Each publisher preferred a different paper. As a consumer, you are the most important judge of quality. Please rank the books in order of image quality.

Two press runs were used to produce the psychophysical sample books. The first press run produced test targets for physical measurements and IT8.7/4 CMYK targets for the profiling process. Files were prepared as was previously discussed.

The experiment stipulated that both a liquid toner and a dry toner press be included. Each press-run included a run on an HP Indigo 7000, a liquid toner press, and a Kodak NexPress S3000, a dry toner press. The presses were carefully calibrated prior to each run. The calibration process was different for each press. PAL uses a NewPage Sterling Ultra Digital stock as the calibration substrate for both presses. Table 6 shows a summary of the calibration procedure for the Kodak NexPress S3000, and Table 7 shows the calibration procedure for the HP Indigo 7000.

Table 6. Calibration procedure for the Kodak NexPress S3000

<table>
<tr><th colspan="2">Kodak NexPress S3000</th></tr>
<tr><td>1</td><td>Color-to-Color Cross Track Calibration: This procedure included a calibration of color registration across the image direction and machine turn registration conducted within the press.</td></tr>
<tr><td>2</td><td>Kodak Intelligent Calibration Solution (ICS): This procedure fixed streaking within tints (due to recently changed parts) and conducted an image-to-sheet registration calibration.</td></tr>
<tr><td>3</td><td>Substrate In-Track Calibration: This procedure conducted a per-substrate registration. This calibration was done independently for each paper, using about 30 sheets per stock. Results were stored in the main press computer.</td></tr>
<tr><td>4</td><td>Color Calibration/Linearization: An X-Rite QuickCal spectrophotometer measured Status T density of patches on a test target produced by the press. Following is a list of densities to which the machine was set.
<table>
<tr><th>Color</th><th>D Meas.</th><th>% Off</th></tr>
<tr><td>Cyan</td><td>1.6</td><td>2.40%</td></tr>
<tr><td>Magenta</td><td>1.66</td><td>1.27%</td></tr>
<tr><td>Yellow</td><td>0.99</td><td>1.39%</td></tr>
<tr><td>Black</td><td>1.63</td><td>0.79%</td></tr>
</table></td></tr>
</table>

Table 7. Calibration procedure for the HP Indigo 7000

<table>
<tr><th colspan="2">HP Indigo 7000</th></tr>
<tr><td>1</td><td>First Transfer Calibration: This procedure adjusted the distance between the pip and imaging blanket.</td></tr>
<tr><td>2</td><td>Full Color Calibration: Paper, ink, and screen ruling information was input into the press. The screen ruling was set to the default 180 HDI (180 lines per inch with High Definition Imaging).</td></tr>
<tr><td>3</td><td>LUT Scenario: The standard LUT for the press was chosen, EXP05 LUT, which uses a 64% output for a 50% tint. Following is a list of target densities to which the machine was set and measured densities after calibration. The tolerance was 0.04D.
<table>
<tr><th>Color</th><th>Target D</th><th>Meas. D</th></tr>
<tr><td>Cyan</td><td>1.75</td><td>175.00%</td></tr>
<tr><td>Magenta</td><td>1.1</td><td>113.00%</td></tr>
<tr><td>Yellow</td><td>1.45</td><td>146.00%</td></tr>
<tr><td>Black</td><td>1.45</td><td>143.00%</td></tr>
</table></td></tr>
</table>

Following calibration, the measurement and profiling test targets were printed on each press. The order in which the papers were run through each press was completely randomized. Duplicates were printed of the test target file for each paper.

IT8.7/4 test targets were printed for each substrate in two different orientations and in four different locations on four separate sheets. The targets were measured using an X-Rite iSis XL and Measure Tool 5.8.10. ProfileMaker 5.8.10 was used to create profiles.

Table 8 shows the settings selected in ProfileMaker to create profiles for each paper and press combination.

Table 8. Settings used to generate profiles for each paper and press combination using Profile Maker 5.8.10

Main Printer Profile Options	
Profile Size	Default
Perceptual Rendering Intent	Paper-colored Gray
Gamut Mapping	LOGO Classic
Viewing Light Source	D50
Separations Options	
CMYK Max	280%
Black Max	100%
Black Start	10%
Separation	MaxK
Define Black Point	CMYK = (67, 57, 56, 100) (Neutralized)
Black Width	100

The second press run occurred one week after the test target and profiling press run. The samples were first scheduled for printing on the HP Indigo 7000, then on the Kodak NexPress S3000. However, the HP Indigo 7000 color management process is different from a conventional color management process as would be used, for example, in offset lithography. The planned color management workflow embedded the ICC profile in each of the four printed images in Adobe Photoshop CS4 by converting the original RGB file to the output profile using the relative colorimetric rendering intent (see Figure 9). Therefore, regardless of the profile embedded when saved, the CMYK values were still set. This is akin to the conventional offset lithography color management process where the ICC profile is maintained through the process. After embedding the profiles in the images through Photoshop, the images were placed into an Adobe InDesign CS4 layout, maintaining the profile embedded in each image, and then saved as a PDF, maintaining the embedded profile here as well. However, after learning about the HP Indigo 7000 workflow, it was determined that this was not the best method. The HP Indigo 7000 RIP interprets ICC profiles embedded within images as ICC simulation space profiles as opposed to ICC output profiles. When the HP Indigo 7000 RIPs a PDF, images within the document are converted to the default output profile stored within the RIP. An alternate workflow was devised as follows. The ICC profiles embedded within the original PDF files were removed using Enfocus PitStop Pro. However, the CMYK values remained the same regardless of the embedded profiles because the CMYK values were retained from the original RGB to CMYK conversion done in Adobe Photoshop CS4. The numbers were verified in Acrobat Professional 9 using Enfocus PitStop Pro.

The HP Indigo 7000 RIP also allows for the loading of custom output profiles. However, it will only convert to this custom output profile if the images contain embedded ICC profiles (which it considers the simulation space). If no profile is embedded, straight CMYK values are printed without converting to the output profile.

The following tests were run to verify these three methods: (1) the file with embedded profiles was run using the press default output profile, (2) the file with embedded profiles was run using the custom-made output profile (the same profile that was embedded in the images), and (3) the file without embedded profiles was printed. Prints (2) and (3) appeared nearly identical while print (1) had a greenish color-cast. This verified that printing the file without an embedded profile and maintaining the embedded profile while loading the profile into the HP Indigo 7000 RIP would produce the same results. The most pragmatic method of the two was to print the files without embedded profiles. This minimized the influence of the RIP on the color management of the final image. The final workflow is shown in . This color management workflow was also used during the Kodak NexPress S3000 press run. While the NexPress handles embedded profiles in a different way than the HP Indigo 7000, bypassing the press's color management system was the simplest approach to minimize color error and maintain the predictability of the process. It is important to note that the methods described here are specific to this experiment's requirements and are not being promoted as a recommended workflow for all cases.

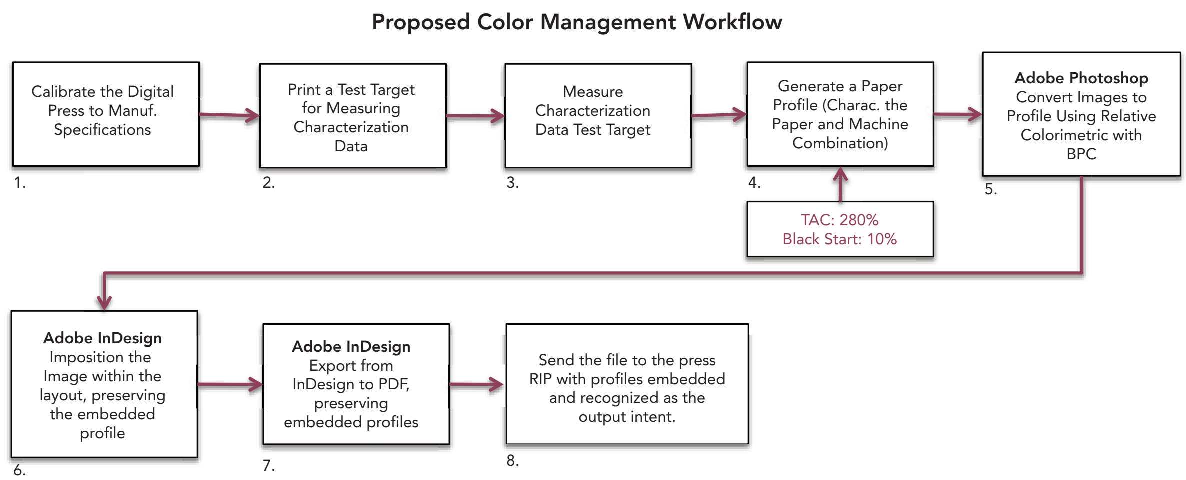


Figure 9. Proposed color management process for producing the sample books during the second press run[10]

10 - The direction of flow is illustrated by the blue arrows.

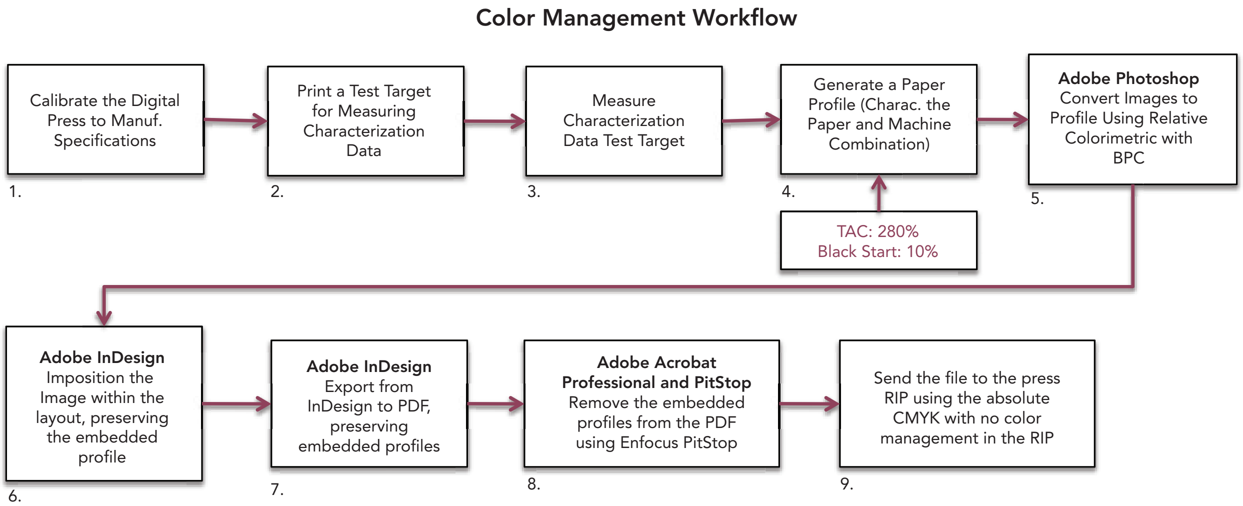


Figure 10. Final color management process used to produce the sample books during the second press run[11]

The profiles for each paper and press combination were evaluated following the second press run. The profile evaluation method was based upon the workflow outlined by Fraser, Murphey, and Bunting (2005), shown in Figure 11. Included in the second press run with the test samples was an IT8.7/4 target converted to the output profile. This target will be referred to as the Profile Evaluation Print (PEP).

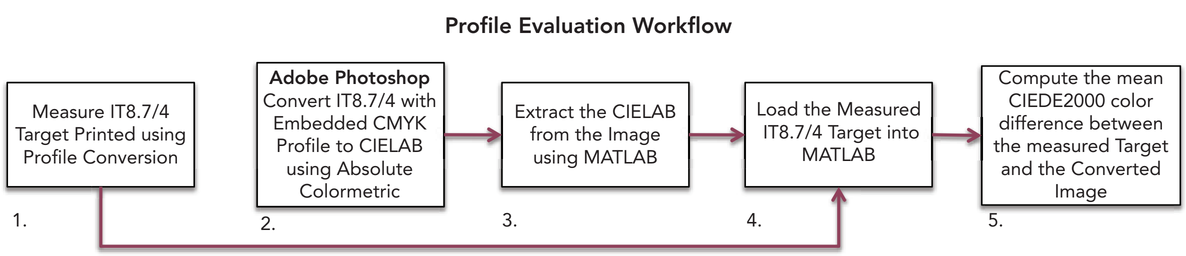


Figure 11. Profile evaluation process used to analyze the success of the color management workflow

The PEP was measured using the X-Rite iSis XL scanning spectrophotometer. A legacy image of the IT8.7/4 was assigned the output profile, then converted to CIELAB using the Absolute Colorimetric rendering intent in Adobe Photoshop CS4. The image was saved then opened in MATLAB®. A MATLAB® function read the CIELAB values from each patch on the target. These CIELAB values served as the reference set. The CIEDE2000 color differences between the measured target and the CIELAB image of the target were calculated. Figure 12 shows the mean CIEDE2000 color differences for each of the 12 papers printed on the Kodak NexPress S3000 and HP Indigo 7000 at PAL.

11 - The direction of flow is illustrated by the blue arrows.

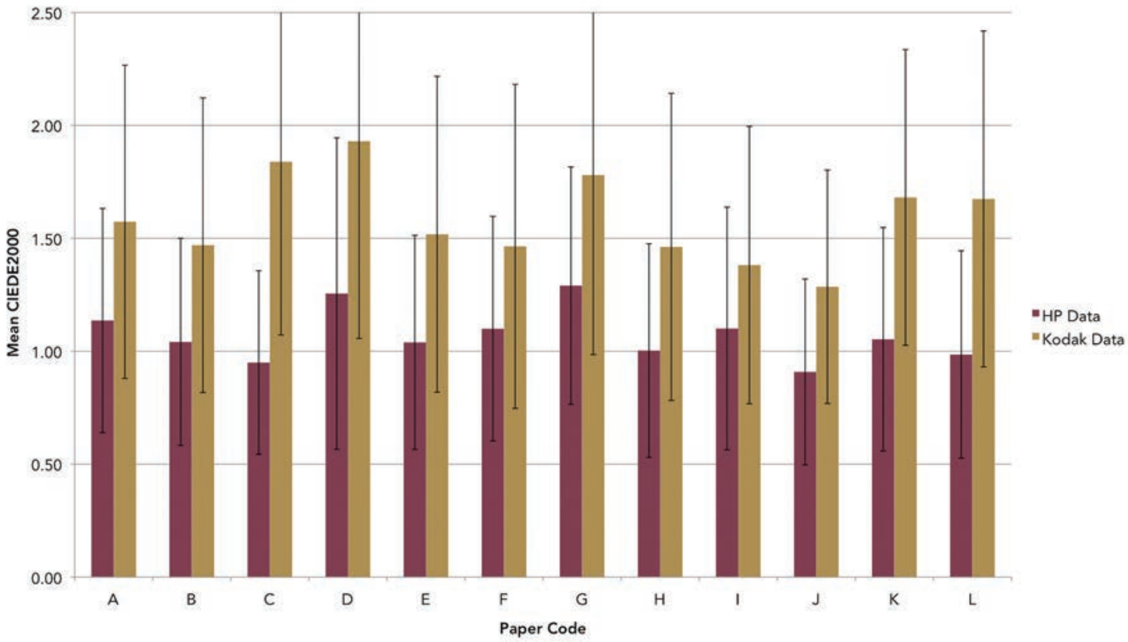


Figure 12. Mean CIEDE2000 error and confidence intervals for the 24 profile evaluations

Psychophysical Data Collection Interface

A MATLAB® interface was designed to collect and analyze data from the psychophysical experiments. Figure 13 shows the six different GUI panels used for data collection. Figure 13a is the introductory component of the interface. The user can either collect data from a new experiment or analyze the collected data. Figure 13b appears when the "Data Collection" option is selected in Figure 13a. The GUI in Figure 13d, summoned when the user clicks the "Collect User Information" button Figure 13b, allows the experimenter to input identification and demographic information about each user, including: experimental order number, press on which the samples were printed for each user name, user initials, user age, user gender, user field and field code, whether the user was experienced with the images, and whether the user was receiving credit for participation. In addition, Figure 13e was used to collect data from the Ishihara Test for Color Blindness.

A file name unique to each observer was constructed from the data from Figure 13d. The file names contained the following elements: order number, press code (0 for the HP Indigo and 1 for the Kodak NexPress), age, field code, experience code (0 for having never seen the images and 1 for having seen the images), class credit code (0 for not receiving credit and 1 for receiving credit), initials, and gender. The following is an example of a file name, "color_17_1_25_3_0_0_RRH_F." This file name indicates the data is from the Color Rendering Quality Experiment, the observer was the 17th to participate, the press number was 1 (Kodak NexPress S3000), the observer was 25 years old, the field code was 3 (Imaging/Color Science or related), the observer was neither experienced nor receiving credit, had the initials RRH, and was female. Further information, such as the observer's full name, e-mail, and full field description, was kept within a separate branch of the data structure and available only to the experimenter.

The rank data for each section of the experiment was collected using the GUI in Figure 13f. Each table allows the experimenter to input the rank order per image. The number next to each table field shows the order in which that image is presented. For example, in Figure 13f, the Oil Painting was presented first, followed by the Aquatint, Watercolor, and the Photograph. The order was determined by a random permutation of the set {1,2,3,4}. Figure 13f also contains two buttons to start and stop a timer. The time started before the observer began ranking the first set of images and ended when the observer finished ranking the fourth set of images.

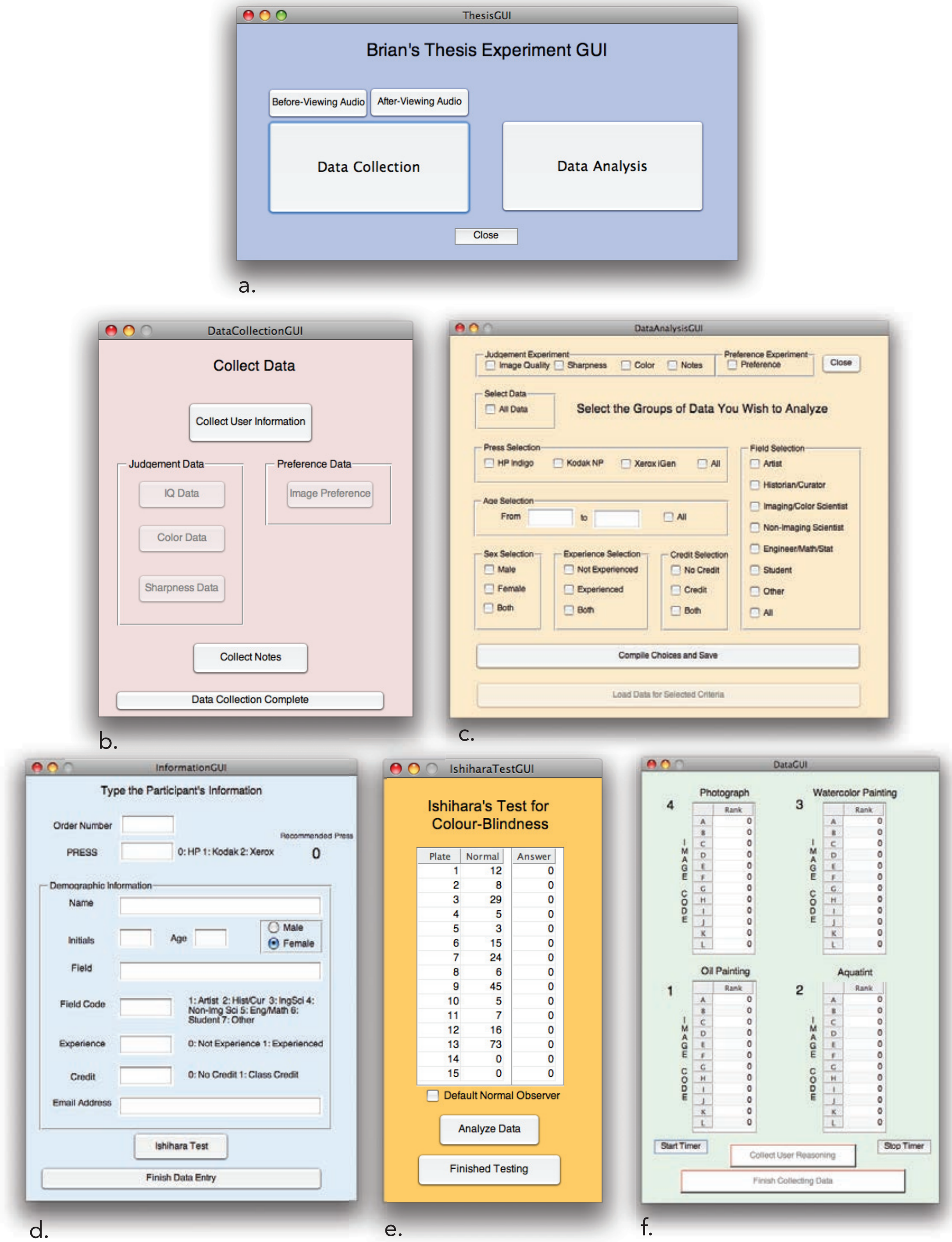


Figure 13. Interface for collecting and analyzing the psychophysical data
Components included: (a) main window, (b) main data collection window with options to run the four experiments, (c) main data analysis window with various analysis data sorting options, (d) observer information window, (e) Ishihara result collection window, and (f) psychophysical data collection window.

Experimental Design

The experiment took place at the Munsell Color Science Laboratory, (MCSL) Building 18 (COL) at the Rochester Institute of Technology, in Rochester, NY. The layout of MCSL is shown in Figure 14.

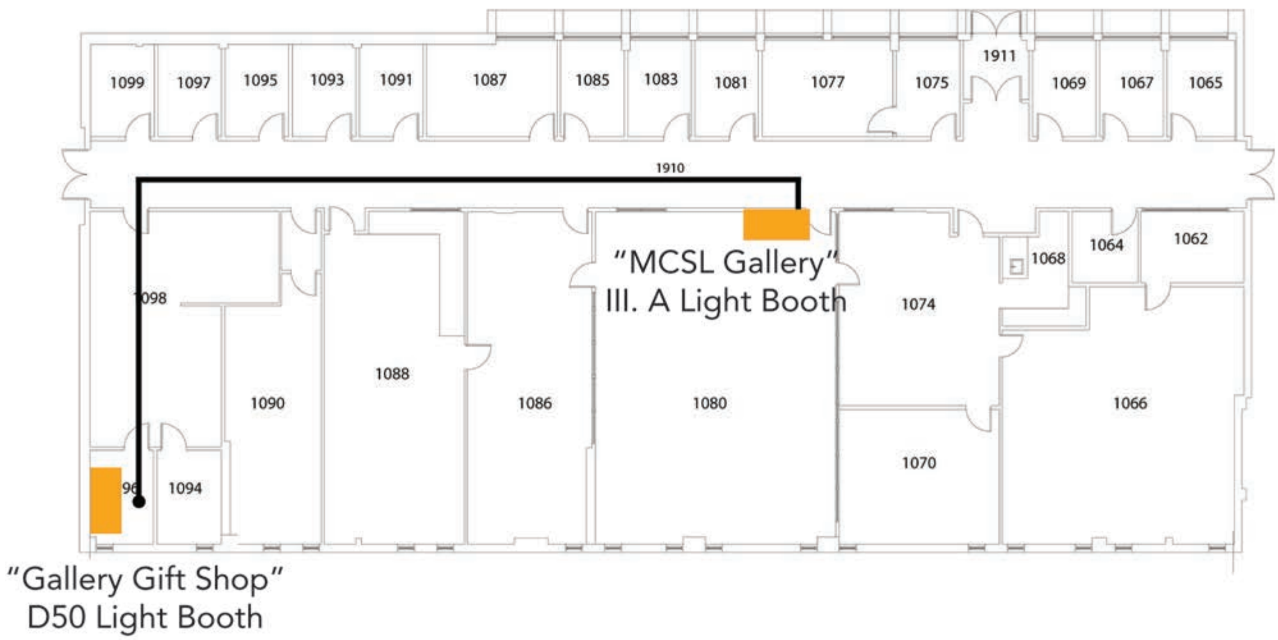


Figure 14. Floor plan of Building 18 at RIT, home of the Munsell Color Science Laboratory[12]

The experimenter met each observer in room 1074. Observers were led to the first component of the experiment, the MCSL Gallery, where they were asked to view original works of art under incandescent illumination in a Macbeth Spectralight II light booth. They were read the following script while standing before the light booth as shown in Figure 15.

> You are the patron of the world-famous Munsell Color Science Laboratory Gallery. Before you are six original works of art in a light booth simulating the art gallery experience. Take a minute or two to become familiar with the artwork, as you would in a real gallery. When you have finished viewing the artwork we will move onto the "Gallery Gift Shop" to conduct the experiment.

The light booth, shown in Figure 15, contained the four works of art used in the experiment with two additional works not used in the experiment. These two additional pieces were added to the light booth to make the experiment more like the experience an observer might have in a real gallery. An observer visiting a gallery, such as in a museum, rarely has the original work present when a patron purchases a book. In

12 - The black line shows the path observers took from the pre-experiment viewing booth to the room in which the experiment was conducted, the "Gallery Gift Shop."

addition, the patron does not know whether all of the pieces on display will be present in a purchased book or which of those pieces will have been printed.

Figure 15. Experimental viewing setup with the Macbeth Spectralight III light booth, illuminated by source 'A,' used to acquaint observers with the original works of art[13]

A brief survey of the lighting in the Rochester Memorial Art Gallery revealed that the majority of the display lighting to be either incandescent or tungsten halogen. Therefore, observers viewed the original artwork under simulated incandescent illumination. The viewing geometry was near to 45/0. The pieces were displayed on a custom-built shelf made of white foam core.

Observers indicated when they were finished viewing the original artwork. The experimenter then led observers to the viewing room, dubbed the "Gallery Gift Shop," where the experiment took place under simulated D50 lighting. The illumination was provided by a bank of fluorescent lamps. Correlated-color-temperature and luminance measurements were made of the illumination in the viewing room from eight positions on the sample display table. The positions of those measurements are shown in Figure 16.

13 - Pieces B, D, E, and F were included in the study. Pieces A and C were used as decoys.

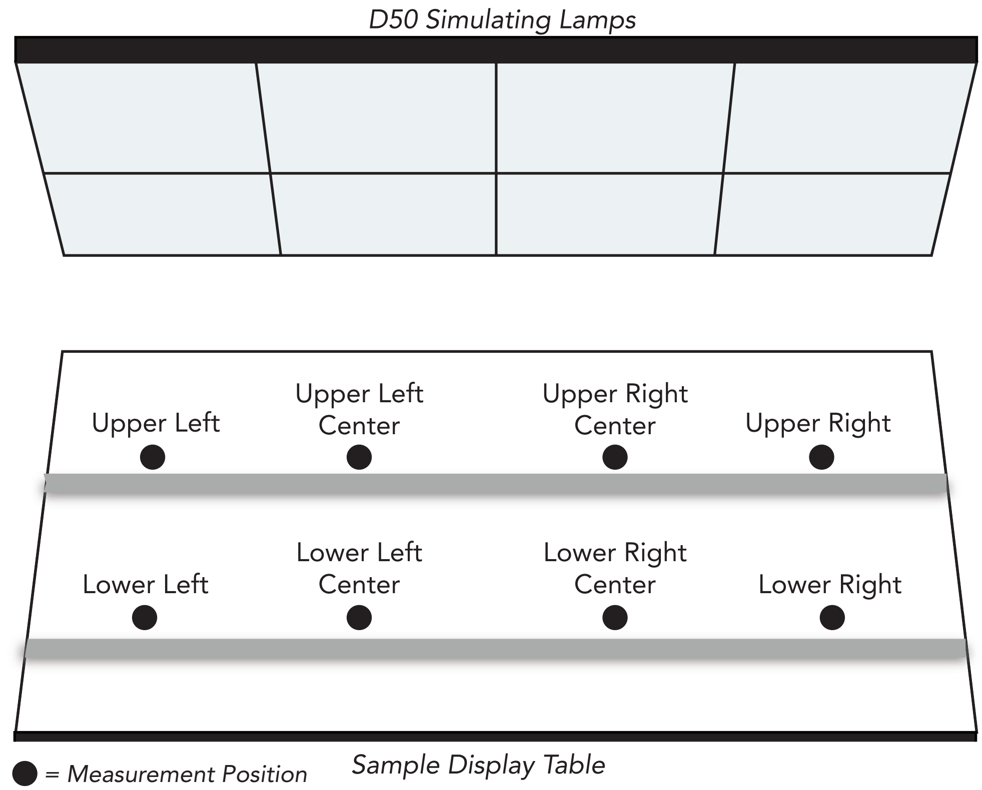


Figure 16. Location of the four fluorescent light panels and the eight measurement positions in the D50 viewing booth

The CCT and luminance measurements made from the eight positions are shown in Figure 17. The plot of CCTs suggests that at least one of the panels exceeded the rated 5000K by 200K. The remaining three panels differed by no more than 50K from 5000K. The luminance of the four panels does seems to increase from right to left. Interestingly, the panel with the highest luminance, the left-most panel, is also the most inaccurate with regards to CCT. Of course, it was expected that the luminance of the bottom tier measurements would be lower than the top due to the increased distance from the light source. This was unavoidable, but was accounted for by randomizing the position of the sample books for each observer. In addition, observers were allowed to move the sample books during the experiment to preferred positions.

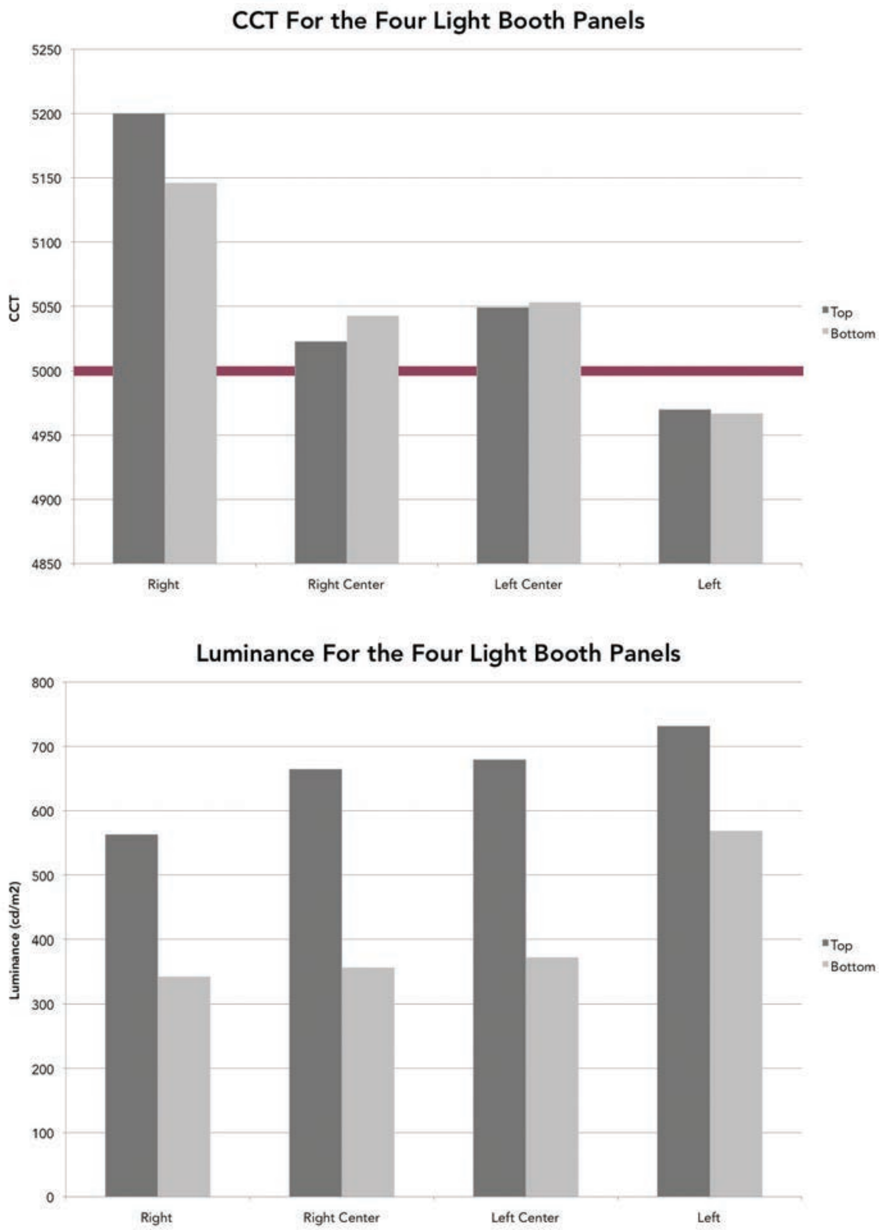


Figure 17. Measurements of CCT and luminance in cd/m² from the eight measurement positions[14]

The experiment was designed in four sections: Image Quality, Color Rendering Quality, Surface Appearance Quality, and Preference. Pre-recorded audio introductions preceded each section. Observers were asked to rank the books based upon the criteria outlined for each sections. Two wood boards (1.5” x 72” x 0.5”) were clamped to a Commando XX tilting table (37” x 72”). The table was tilted to about 25 degrees off the vertical. The boards were spaced such that the sample books could comfortably be placed on the bottom and top boards. A schematic of the viewing environment is shown in Figure 18.

14 - The plots compare measurements made from the top tier and bottom tier on the sample table. The red line in the CCT chart shows the position of 5000K, the rated CCT for the fluorescent lamps.

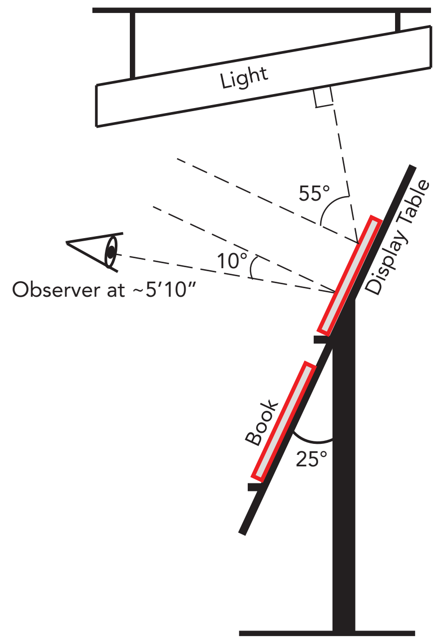


Figure 18. Viewing and lighting geometries experienced by observers during the psychophysical experiments[15]

Each board could hold six to seven books (see Figures 19a and c). All four images were ranked independently in the Image Quality, Color Rendering Quality, and Surface Appearance Quality sections. The following instructions were provided to each observer for the first three experiment sections:

Image Quality:

> You are now in the gallery gift shop and intend to purchase a book containing reproduction of those six pieces. However, the available books contain reproductions of only four of the six pieces. Twelve publishers produced the same book. Each publisher preferred a different paper. As a consumer, you are the most important judge of quality. Please rank the books in order of image quality. You will rank each image separately. I encourage you to handle the books to gain the full experience. However, please do not flip to the next image.

Color Rendering Quality:

> You will now rank the books on the basis of color rendering. Please rank the books based upon the following criterion: HIGHEST QUALITY COLOR RENDERING. Your criteria for ranking highest quality color may only contain factors relating to the color of the image.

15 - The dotted lines show the directions of light and viewing. The illumination angle and average observer viewing angles are also shown.

Surface Appearance Quality:

> You will now rank the books on the basis of surface appearance. Please rank the books based upon the following criterion: HIGHEST QUALITY SURFACE APPEARANCE. Your criteria for ranking highest quality surface appearance may contain factors such as texture and gloss. However, you must completely disregard color when making your judgments.

Observers were instructed to use the provided space however they felt was most efficient for making their judgments, and were allowed to handle the books as they saw fit. All observers wore white cotton gloves while conducting the experiment. Observers handed books to the experimenter in the order of image quality, either from best to worst or worst to best. The experimenter recorded this order. However, beyond that method, observers varied greatly in how they used the space to make their judgments. The most common technique observers used was to arrange the books in order of quality on the table before delivering them to the experimenter. Some observers simply handed the books to the experimenter one-by-one as they decided upon the rank. Figure 19b shows an observer conducting the experiment.

a.

b.

c.

Figure 19. Experimental display setup

The 12 sample books were positioned on the viewing table (a, c) under the D50 fluorescent lighting. Observers viewed and moved the books around the table as they made their ranking decisions (b).

The final stage in the experiment was the judgment of Preference. The experimenter placed all 12 books on the table with the Oil Painting facing outward because it was the first image in the book. The observer's task was to choose their top three favorite books. The following scenario was presented to the observers during the audio introduction to this section.

> The books are all priced the same. You now decide to purchase your favorite book. Please select your three most preferred books, and present them to me in the order of most favorite to least favorite. These choices are not based upon your personal feelings. You can use any criteria you wish to make your decision.

Observers were allowed to look at any image and use any criteria they felt important to choosing their most preferred books. Observers handed their top three books to the experimenter after making their decisions.

In addition to the standard instructions given to each observer, a sheet was provided containing suggestions of factors the observers could consider during each of the four experiments. Table 9 lists the suggestions provided to observers during the experiments.

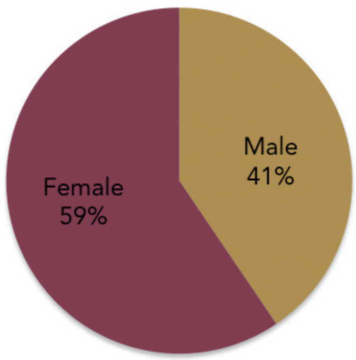
Table 9. Suggestions provided to observers of potential factors to consider

Experiment	Suggestion
Image Quality	Use any criteria you think is important to image quality
Color Rendering Quality	May Include: color cast, overall color, color accuracy, brightness, gloss, contrast, lightness, saturation, etc.
Surface Appearance Quality	May Include: texture, relationship between texture and the image, gloss, evenness of the Ink across the image, opacity of the paper/image, etc
Customer Perceptions	May Include: Your personal feelings, emotions, preferences, etc.

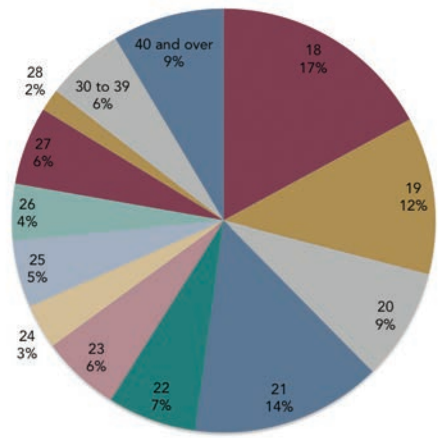
Experiment Observers

One hundred and seventeen observers participated in the Image Quality Experiment. Figure 20 describes the gender, age and field distribution of observers. Sixty-one percent of observers were female and 39% were male (Figure 20a). A wide range of ages were represented; however, 50% of observers were between 18 and 22 years old (Figure 20b). While this age range was representative of the undergraduate student population, not all undergraduate students were between 18 and 22 years old. Some students were in the last year of a five-year program, while others simply began their undergraduate careers when they were older.

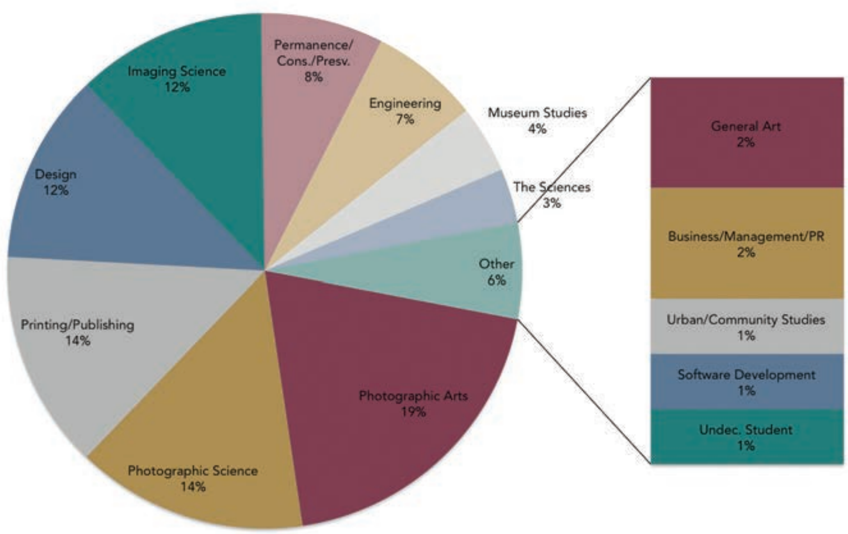
A similarly wide range of fields were represented in the observer population. The fields listed in Figure 20c include undergraduate and graduate fields of study in addition to career fields because such a large percentage of students participated in the study (91% students). Of those students who participated, 65% received extra credit as an incentive to participate. Twenty-nine percent of observers were 18 or 19 years old. The large turnout of observers within this age range is due to the offering of extra credit by a professor of a large freshman undergraduate lecture class. An additional 8% of students were directly affiliated with the RIT Center for Imaging Science.



a.



b.



c.

Figure 20. Observer demographics by (a) gender, (b) age, and (c) field of work or study

All observers were subject to the Ishihara Test for Color Blindness. One male observer was determined to have a red-green color deficiency. The extent of his deficiency was not analyzed beyond the Ishihara Test. However, the color-deficient participant indicated that he had no trouble differentiating between colors within the four images and that he would have no trouble completing the experiment. The most significant color differences were along the yellow-blue opponent color axis due to the differences in paper color. This observer was allowed to conduct the experiment and his data was included with the general population.

Coolness Estimation Experiment

The coolness estimation experiment arose from a decision made shortly after the psychophysical phase of this thesis began. The sample papers were selected using a two-level factorial design. The original plan was to analyze the rank data using that design. However, the two-level design minimized the amount of information available for model construction. Physical measurements of the virgin paper stocks were concurrently taking place and included measurements of roughness, basis weight, gloss, and color, among others. The four factors used to select the samples were basis weight, roughness, gloss, and color. It was simple to include the physical measurements of basis weight, gloss, and roughness in the experimental analysis. However, there was no simple solution available for measuring color on the basis of warmness and coolness, the two levels of color used in sample selection. ISO Brightness was one option considered for color [ISO 2470-2:2008]. However, ISO Brightness was designed for controlling the quality of the bleaching process for non-fluorescent papers (Axiphos GmbH, 2001). Its current use in marketing the brightness of fluorescent papers is an incorrect use of the metric. While it serves its purpose in marketing, its use as a metric for coolness cannot be justified. Similarly, CIE Whiteness (CIE, 2004) could be used as a measure of coolness. However, CIE Whiteness can only be used for white papers under CIE Illuminant D65 and is not appropriate for evaluating papers much departed from neutral or differing in fluorescence, properties held by papers included in this experiment.

A psychophysical experiment was designed to create a quantitative measure of the perceived coolness of white paper. Coolness and warmness are common terms used by members of the art field for categorizing papers. This experiment was designed to develop a colorimetric approach to defining coolness. The resulting coolness metric became the color factor in the experimental data analysis. Following is a description of this experiment.

Twenty samples were included in the experiment, plus an anchor sample. The experiment was designed using the magnitude estimation method (Besore, 1973, pp. 1023–1132). Among the 20 samples, 12 were the sample papers included in the thesis experiment. Eight additional digital press papers, available in the Munsell Color Science Laboratory were included. Table 10 lists the eight additional papers and the anchor sample, along with their associated experiment codes. The 'S' in the code stands for 'supplement.'

Table 10. Papers used in the coolness estimation experiment with experiment codes

Paper Code	Papers
S1	Mohawk Via Satin Bright White 80lb text
S2	Neenah Finch 80lb text
S3 (ANCHOR)	Hahnemuehle Natural Art Duo 256 gsm
S4	NewPage Sterling Ultra Digital 80lb cover
S5	Neenah Classic Crest Text Solar White 80lb Smooth
S6	Neenah Classic Crest Text Classic Natural White 80lb Smooth
S7	Neenah Classic Crest Text Avon Brilliant White 80lb Smooth
S8	Neeah Classic Crest Text Recycled 100 Bright White 80lb Smooth
S9	Neenah Classic Crest Text Avalanch White 80lb Smooth

Each sample measured 8 in^2. A 2 in^2 patch of white paper was visible in the middle of each sample surrounded by near 50% gray paper. The samples were mounted on black foam core. Figure 21a shows an array of the samples laid out on the experiment table.

Figure 21. Coolness estimation experiment setup

The samples were laid out on the display table (a) before beginning the experiment. Observers viewed each sample with the anchor (b, c) such that the samples subtended a two-degree field of view.

Observers were instructed to stand on a black line marked 1.2 meters from the samples such that the white patches on each sample subtended a two-degree visual field. Samples were shown in a random order, always paired with the anchor. The anchor sample was chosen for its relatively flat spectral curve and lack of optical brightening agents. The anchor sample was assigned an arbitrary coolness value of 100. Observers were instructed to estimate the magnitude of the sample's coolness compared to the anchor sample. Following are the instructions given to observers:

> Coolness is here defined as the sensation that a paper appears cool colored. It is counter to warmness, the sensation that a paper appears warm colored. In this experiment, you will measure the coolness of a series of paper samples. The coolness of the anchor sample on the board is 100. You will be presented with samples of white paper similar to the anchor sample. Your task is to estimate the coolness of the white paper samples relative to the anchor sample. You may estimate the coolness as whatever it appears to you, but you should maintain the appropriate ratio relationships between your estimates. For example, if you are presented with a sample having half the coolness of the anchor, you should call it 50; and if it is twice as cool, you should call it 200. I recommend you make your judgments as quick as possible. Remember, the anchor sample has a coolness of 100.

Figure 21b shows an observer participating in the experiment and Figure 21c shows an example of a sample pair shown to each observer. On the left of Figure 21c is the anchor sample and on the right is a test sample.

Forty observers participated in the Coolness Estimation experiment. Eighteen observers were male and 22 observers were female. The average age of observers was 25, with 18 as the age of the youngest observer and 49 as the age of the eldest. All observers had normal color vision, as determined by the Ishihara Test for Color Blindness.

Magnitude estimation using a single anchor allows for the generation of a ratio scale. However, because minimal constraints were placed on observer responses, the direct estimates made by each observer could not be used. For example, the largest coolness estimate for one observer may have been 150, while, for another, it may have been 300. The estimated coolness for these two observers cannot be compared unless they are normalized to the same scale (while maintaining the ratio relationship between all estimates for each observer). Therefore, the following procedure was used to calculate the Normalized Scale Position (NSP) of observer responses based upon methods outlined by Engeldrum (2001) and the *SPIE Handbook of Photographic Science and Engineering* (Besore, 1973).

An assumption was made that observer responses followed a log-normal distribution (Engeldrum, 2001). The first step was to normalize the magnitude estimates of each observer such that they lay on the same scale. Equation 3 was used to scale observer estimates. The difference between the grand mean of the log estimates and the column means of the log estimates was added to the log estimates in each row. This centered each observer's estimates along the population mean, M_{in}.

$$\mathbf{M}_{in} = exp\left(ln(\mathbf{R}_{in}) + \left[\frac{1}{IN} \sum_{i=1}^{I} \sum_{n=1}^{N} ln(\mathbf{R}_{in}) - \frac{1}{I} \sum_{i=1}^{I} ln(\mathbf{R}_{in}) \right] \right) \tag{3}$$

where I is the number of papers and N is the number of observers. The geometric mean was calculated for each column respective to the 12 images as shown in Equation 4.

$$\mathbf{M}_n = \left(\prod_{n=1}^{N} \mathbf{R}_{in} \right)^{\frac{1}{N}} \tag{4}$$

The ratios between pairs, Equation 5, were calculated using mean estimates from Equation 4. The ratio matrix, ***R***, is a 12x12 matrix with rows and columns corresponding to the 12 images.

$$\mathbf{R} = \begin{bmatrix} 0 & \frac{M_2}{M_1} & \cdots & \frac{M_{12}}{M_1} \\ \frac{M_1}{M_2} & 0 & \cdots & \frac{M_{12}}{M_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{M_1}{M_{12}} & \frac{M_2}{M_{12}} & \cdots & 0 \end{bmatrix} \tag{5}$$

The column sums of **R** were calculated using Equation 6.

$$R_{i,sum} = \sum_{j=1}^{J} \mathbf{R}_{ij} \tag{6}$$

The column sums of R were used to sort the columns of R in descending order. The presorted columns sums were transposed to the rows of R such the rows could be sorted on the same basis as the columns. The resulting matrix $\boldsymbol{R}_{sorted}$ had a diagonal of zeros, verifying that the structure of the matrix remained unchanged from R. Figure 22 illustrates the transposition of column sums to the rows.

$$\mathbf{R}_{sorted} = \begin{bmatrix} 0 & \frac{M_2}{M_1} & \cdots & \frac{M_{12}}{M_1} \\ \frac{M_1}{M_2} & 0 & \cdots & \frac{M_{12}}{M_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{M_1}{M_{12}} & \frac{M_2}{M_{12}} & \cdots & 0 \end{bmatrix} \begin{matrix} R_{1,sum} \\ R_{2,sum} \\ \vdots \\ R_{12,sum} \end{matrix}$$
$$\begin{matrix} R_{1,sum} & R_{2,sum} & \cdots & R_{12,sum} \end{matrix}$$

Figure 22. Transposition of column sums to the rows for the R_{sorted} matrix.

Column ratios were calculated using Equation 7.

$$\begin{aligned} \mathbf{CR}_1 &= \mathbf{R}_{(sorted,2,1)} \\ \mathbf{CR}_{i=2:11} &= \frac{\mathbf{R}_{(sorted,1,i)}}{\mathbf{R}_{(sorted,1,i+1)}} \end{aligned} \tag{7}$$

The calculation of column ratios is based upon the principle shown in Equation 8. The mean of the ratio estimates in Figure 22 is the best estimate of the ratio R_1/R_2 (Besore, 1973). However, after analyzing the data in this experiment, it was shown that the ratios in Figure 22 were identical. Thus, the column ratios, Equation 7, were calculated using a single row in $\mathbf{R}_{sorted}$. The first element **CR** is the first element in the second row of $\mathbf{R}_{sorted}$. The remaining elements of **CR** are the ratios between adjacent elements of the first row in $\mathbf{R}_{sorted}$.

$$MSP = \frac{1}{I-1}\sum_{i=1}^{I-1}\mathbf{SP}_i \tag{8}$$

Scale values were calculated from the column ratios. The value, 10, was chosen as an arbitrary scale position for the 12th element of the scale position array, ***SP***. This value marked the psychological distance between the sample with the lowest estimated coolness and the zero position. Remaining elements of **SP** were determined using Equation 9 as the product between the i^{th} **CR** and the i^{th} +1 **SP**.

$$\mathbf{NSP}_i = 100 * \frac{\mathbf{SP}_i}{MSP} \tag{9}$$

The mean scale position, ***MSP***, was calculated using Equation 10.

$$\frac{R_1}{R_2} = \frac{R_1/R_3}{R_2/R_3} = \frac{R_1/R_4}{R_2/R_4} = \frac{R_1/R_5}{R_2/R_5} \tag{10}$$

The MSP was used to normalize the scale positions to 1. The final ***NSP*** values were calculated using Equation 11. The mean was set at 100.

$$\begin{aligned}&\mathbf{SP}_{12} = 10\\&\mathbf{SP}_{i,(11:-1:1)} = \mathbf{CR}_i * \mathbf{SP}_{i+1}\end{aligned} \tag{11}$$

Physical Measurements

While the psychophysical experiments provided all of the information necessary to model Customer Perceptions and Customer Image Quality Rating, the Image Quality Circle (Engeldrum, 2004b) was not complete without the measurement of Physical Image Parameters (PIPs). The papers used in the psychophysical experiment were selected on the basis of color, basis weight, roughness and gloss. Physical measurements of the papers were made of these four factors to satisfy the Physical Image Parameter section of the Image Quality Circle. In addition, several other factors not included in the paper selection process were also measured. Engeldrum used the term "Visual Algorithms" to describe the modeling of Customer Perceptions by Physical Image Parameters. Physical measurements of coolness, basis weight, roughness and gloss were used to develop a base visual algorithm to predict Color Rendering Quality and Surface Appearance Quality. However, a more detailed model containing Physical Image Parameters not included in sample selection was also developed. The following sections describe the physical measurement procedure for the sample selection parameters and the additional parameters.

Measurement of Basis Weight, Roughness, and Gloss

The procedure for measuring basis weight and grammage is standardized by TAPPI Test Method T410 om-98. The mass of ten 20 cm by 25 cm sheets was measured for each paper. The mass of each stack was divided by 10 to calculate the mass of a single sheet. Basis weight is calculated from grammage. Grammage is measured in grams per square meter. The mass of each sheet was divided by the area of the sheet in square meters to calculate grammage. TAPPI T410 om-98 defines different grammage to basis weight conversion factors for different paper classifications. Basis weight is physically measured by weighing a stack of 500 sheets of paper cut to specific dimensions. Each paper classification, such as bond, manuscript cover, blotting, cover, tissue, and newsprint, has specified dimensions to which a stack of 500 sheets must be cut. The papers used in this experiment were classified as "Book" paper. The conversion factor between grammage and basis weight for Book paper is 0.676. The calculated grammage of each paper was multiplied by the conversion factor to calculate basis weight.

Roughness was measured using the ISO 8791-4:2007 (ISO 8791-4, 2007), Print Surf method. A Testing Machines Inc. Parker Print Surf device was used to measure roughness. Five roughness measurements, measured in micrometers of air flow across the paper surface, were made of each paper, and averaged for the final measurement.

Gloss was measured using the 60-degree method, as described by ASTM D523-08 (ASTM D523, 2008) and ISO 2813:1994 (ISO 2813, 1994). A Color Control Systems ETB-0833 Glossmeter was used to make five measurements of each sheet along the machine direction. The five measurements were averaged for the final measurement.

QEA Image Analysis Tool Measurements

The IAS Test Targets were scanned using an Epson Expression 10000XL flat-bed scanner. IASLab™ allows for an automated sequence of measurements from a single scanned target. Figure 23 shows the layout of the automated measurements.

IASLab™ has several different tools available for print analysis. Each tool provides multiple types of data. An area analysis was performed on the 40% and 100% CMYK tone reproduction solid ink patches from which the mottle data was collected. The IAS Test Target contained Dot Quality images with 3x4 arrangements of 0.1 mm through 0.6 mm dots (see the Dot Quality section in Figure 23). A dot analysis was performed on all six dot sizes for cyan, magenta, and black. Yellow was not included because the software had difficulty detecting the dots due to the low contrast between the yellow toner and the paper. The mean dot circularity was collected from each of the dot analysis measurements. The third set of measurements used the section of the IAS Test Target (see the top-most set of images in Figure 23) designated for Line Quality, Width, Density, Raggedness and Blurriness. Each CMYK image consists of horizontally and vertically arranged colored lines imposed on white and white lines imposed on colored patches. Each array of lines stepped in thickness from 2 pt to 1/4 pt in 1/4 pt increments, and also included a 1/8 pt line. A line analysis was performed on the colored lines imposed on white for both horizontal and vertical arrangements. Lead and tail line raggedness measurements were recorded from each set of measurements provided by the line analysis.

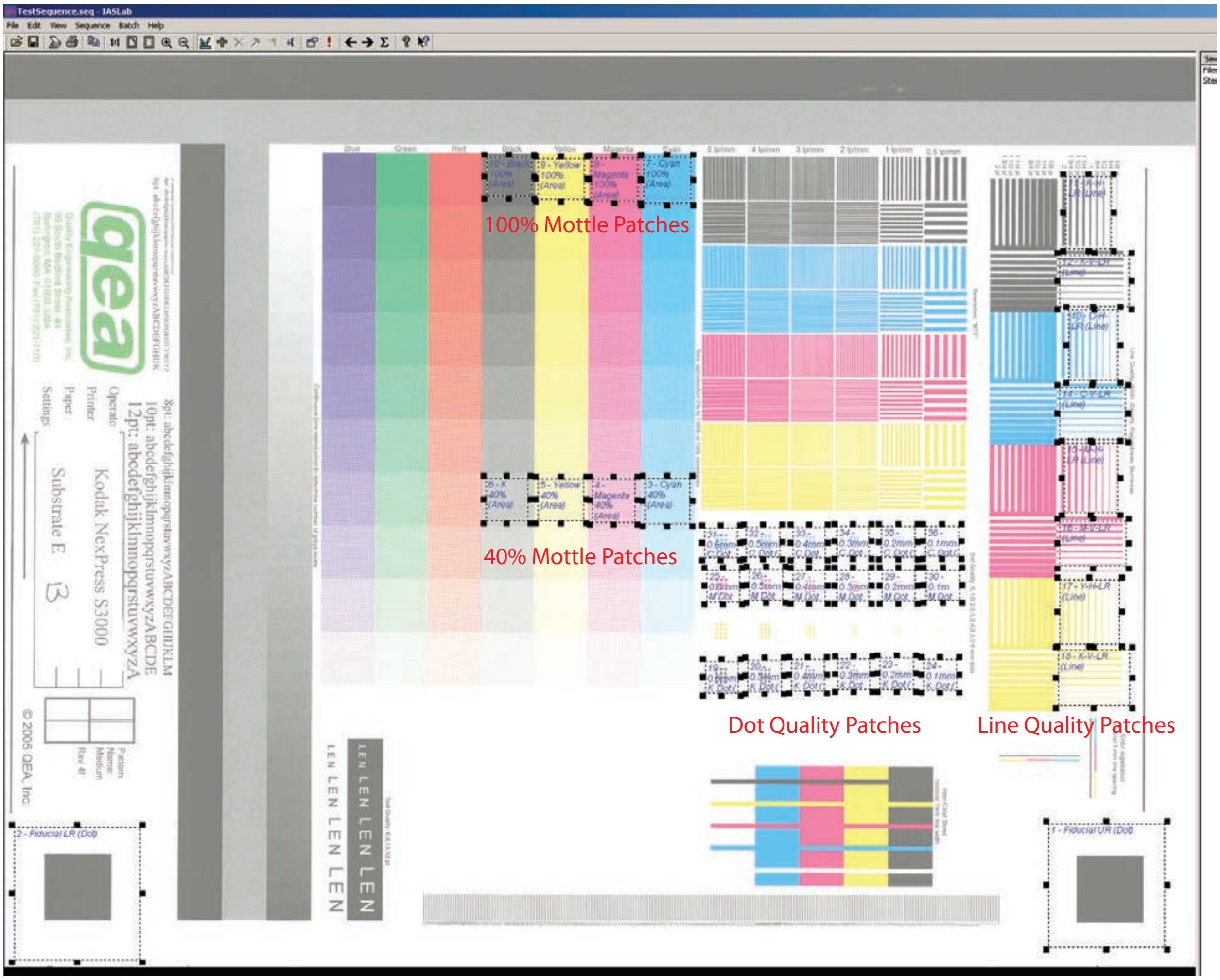


Figure 23. Screen capture from IASLab™ showing the layout of automated measurements

Additional Measurements

Print gloss was measured using the same method as for the measurement of paper gloss. Sixty-degree gloss was measured on the 100% CMY patch from the target shown in Figure 7. The method, described in ISO 19799:2007 and ASTM D7163-05, does not specify a specific solid area for measurement and allows the use of 20, 60, and 85 degree geometries.

Caliper was measured for each paper stock using the standard methods. All caliper measurements, recorded in units of micrometers, were conducted at the PAL Print Materials and Analysis Lab.

It was previously discussed how opacity and basis weight were confounding factors in sample selection. Opacity, standardized in TAPPI T425 om-01, was measured for the 12 papers using a Technidyne Technibrite Micro TB-1C. The average of five measurements was recorded for each paper.

While opacity is a standard metric, it is only a measurement of the light reflected off a paper set over a black-trap. It does not account for ink-penetration into the paper surface. The International Paper Company developed an additional metric for the measurement of Print Show-Through (J. Kohler, personal communication, January, 2011). PST is not a direct measurement of opacity, as standardized in TAPPI T425 om-01. First, ΔE^*_{ab} is measured between virgin white paper and the backside of a black solid area print. The International Paper method used a 270% black solid area patch. However, such a patch was not printed in this experiment. A 300% CMY patch was used instead. Second, ΔE^*_{ab} between virgin white paper and a black solid area print (directly on the print, as opposed to a backside measurement) was measured. The percent PST is the ratio between the backside ΔE^*_{ab} and front side ΔE^*_{ab} (see Equation 12).

$$Percent\ PST = \frac{\Delta E^*_{ab,\ backside}}{\Delta E^*_{ab,\ frontside}} \tag{12}$$

The International Paper PST method is a viable complement to opacity. However, their use of ΔE^*_{ab} was questionable. Color information introduces unnecessary error into a measurement that is largely based on density and luminance. Highly fluorescent papers and highly colored papers could result in an incorrect quantification of the visible PST. Color is not important in determining the visibility of an image on the backside of a page. Therefore, the International Paper PST method was modified to include ΔL^*, instead of ΔE^*_{ab}. The metric for percent PST, shown in Equation 13, was included in this research.

$$Percent\ PST = \frac{\Delta L^*_{backside}}{\Delta L^*_{frontside}} \tag{13}$$

The percent PST metric is the best available metric relatable to the visual experience of perceiving backside-printed text on samples during the experiment.

Results and Analysis

Interview Analysis

Question Response Analysis

The interview questions shown in Table 3 are reiterated in Table 11 for convenience. It was previously mentioned that the default interview questions had to be adjusted for each interviewee. The following are examples of such adjustments in addition to analysis of the question responses.

Table 11. Targeted interview questions

<table>
<tr><th>Type</th><th>No.</th><th>Question</th><th>Answer Choices (If applicable)</th></tr>
<tr><td rowspan="12">Demographic Questions</td><td>1</td><td colspan="2">Would you classify your company or field as publishing, printing, an artist, or a user of reproduced fine art?</td></tr>
<tr><td>2</td><td colspan="2">Please describe your position in your company or (if an artist) the medium in which you most frequently work.</td></tr>
<tr><td>3</td><td colspan="2">How many years has your firm (or yourself) been active?</td></tr>
<tr><td>4</td><td colspan="2">How many employees are currently in your company?</td></tr>
<tr><td rowspan="8">5</td><td rowspan="8">Which of the following best describes your company’s 2004 revenues?</td><td>< $3 million</td></tr>
<tr><td>$3 million to $5 million</td></tr>
<tr><td>> $5 million to $10 million</td></tr>
<tr><td>> $10 million to $15 million</td></tr>
<tr><td>> $15 million to $20 million</td></tr>
<tr><td>> $20 million</td></tr>
<tr><td>Don’t know</td></tr>
<tr><td>Refused to answer</td></tr>
</table>

Type	No.	Question	Answer Choices (If applicable)
Direct Research Questions	6	Which digital presses do you use most often for books of reproduced fine art? If you do not use digital presses, which offset presses do you use?	
	7	Which properties of unprinted paper are important to your decision to select a particular paper for use in books of reproduced fine art? Please rate each property on a scale from 1 to 4, where 4 is "critically important" and 1 is "not important."	Grain Direction
			Paper Gloss
			Caliper
			Basis Weight
			Whiteness
			ISO Brightness
			Opacity
			Stiffness
			Other
	8	Which properties of printed paper are important to your decision to select a particular paper for use in books of reproduced fine art? Please rate each property on a scale from 1 to 4, where 4 is "critically important" and 1 is "not important."	Neutral Color
			Print Contrast
			Tone Reproduction
			Solid Ink Densities
			Print Gloss
			Differential Gloss
			Line Quality
			Other
	9	What are the paper brands (manufacturers) you use most frequently?	
	10	From the brand you use the most, which paper grades do you most frequently purchase for non-fine-art reproduction jobs?	
	11	From the brand you use the most, which paper grades do you most frequently purchase for fine-art reproduction jobs?	
	12	Of those papers you listed in the previous question, which do you use most frequently for fine art reproduction?	
	13	Who commonly decides which paper to select for printing the book of reproduced fine art? Please indicate the influence that each group of people has in the purchasing decision on a scale from 1 to 4, where 4 is "all of the weight" and 1 is "none of the weight."	Printer
			Publisher/Curator
			Artist (if living)
			User/ Expert Consultant (if artist is not living)
	14	What role do the following people bring to the process of deciding which paper to use in a book of reproduced fine art? Please indicate the influence that each group of people has in the purchasing decision on a scale from 1 to 4, where 4 is "all of the weight" and 1 is "none of the weight."	Printer
			Publisher/Curator
			Artist (if living)
			User/ Expert Consultant (if artist is not living)
Open Ended	15	Please explain what influence the original artwork's medium has on the choice of paper for reproduction.	
	16	Please explain how the integrity of the reproduction is influenced by the choice of paper.	

It was often clear upon entering into the interviews whether the interviewee was a publisher, printer, artist or user.[16] However, it was also clear that this list was not nearly inclusive enough to cover the range of fields represented by the interviewees. One graphic designer was included as well as a paper manufacturer, and interviewees often played multiple roles. The curators, who commonly produce more small publications than large books, often acted as publishers separate from their jobs as curators and scholars. Thus, Question (Q) 1 did not provide enough information for analysis.

The interviewees' job descriptions were, based on responses to Q2, often either too broad to narrow down to a single sentence, were well described by their title, or were too specific to protect the anonymity of the interviewees. A short summary of interviewee job descriptions can be found in Table 2. However, this is the extent of information provided for Q2.

Questions 3-5 asked for specific demographic information from the interviewees. Unfortunately, these three questions, the responses of which are summarized in Table 12, provided little useful information for this thesis with the exception of demonstrating the variety among institutions and the vast experience represented by the interviewees. Question 3 lacked the specificity to gauge the true level of an interviewee's experience. Some responses regarded the interviewee's active years over a career, while others regarded the interviewee's active years at a particular institution. Question 4 could also have been refined. Some larger institutions employed hundreds of people, while only a few worked with publications. Some institutions had departments working in cooperation where publications were created with input from many different people, often including contractors.

Table 12. Interviewee responses to demographic questions

Interviewee	3. How many years has your firm/yourself been active?	4 . How many employees are currently in your company?	5. What best describes your company's 2009 revenues?
Cur 1	6 in current position, 35 in the curatorial field	10 people in department	$7-9 million institution budget, zero budget for publications
Cur 2	Worked 8 years as a curator in previous position.	Had a staff of 5 to 6	$7-9 million institution budget
Cur 3	Institution has been publishing for 100 years.	Several departments involved in publishing.	$4-4.5 million budget
Pub 1	9 years in current position	20 employees in a photography studio	N/A
Pub 2	20 years in current position	12 people in department	N/A
Pub 3	27 years in current position	39 to 40 people in department	$3 million in revenues
Pub 4	15 years in the field	2 in department, but many in institution	$7-9 million institution budget

16 - No users were interviewed during this research.

Interviewee	3. How many years has your firm/yourself been active?	4 . How many employees are currently in your company?	5. What best describes your company's 2009 revenues?
Pub 5	30 years in current position	20 people in department	less than $3 million
Pub 6	8 years in current position	1.5 full-time plus students	less than $3 million
Pr 1	Company active for 65 years	62 employees at company	$9-12 million in revenues
Pr 2	1 year in current position (newly created subsidiary)	2 employees in subsidiary	less than $3 million
GD 1	5 years in current position, 25 years in the field	2 employees in company	N/A
PM 1	Company producing paper for 75 years	700 employees in company	N/A

Questions 6, 9, 10, 11, and 12, the responses of which are shown in Table 13, provided more useful information than Questions 3-5. Question 6 revealed that, while many interviewees had not yet explored digital printing, those that had primarily used the HP Indigo, the brand routinely cited for its use of a liquid toner system. Questions 9-12 were designed to aid in sample selection for the psychophysical experiment. The goal of sample selection was to include a statistically designed variety of papers while incorporating papers currently used in the industry. Mohawk was mentioned by many interviewees as a well-respected, though expensive, manufacturer of paper for digital presses. PUB 4, PUB 6, PR 1, and GD 1 all mentioned using Mohawk brand papers. PUB 5 mentioned using Sappi McCoy, a line with similar options as Burgo Chorus Art, mentioned by CUR 2 and PR 1. Burgo Chorus Art Gloss Text 80lb and 100lb papers were included in the experiment over the Sappi McCoy because it was felt that they were slightly cooler, providing more color separation than the other possible samples. In addition, Burgo Chorus Art was used in the book published by CUR 3. Although CUR 3 varnished the paper, it was felt that including a brand in the experiment that had been used in a locally printed book gave it greater value than using Sappi McCoy, of which no fine art reproduction samples were available.

Table 13. Interviewee responses to direct research questions

Interviewee	6. Which digital presses do you use most often?	9. What are the paper brands (manufacturers) you use most frequently?	10. Which paper grades do you most frequently purchase for non-fine-art reproduction jobs?	11. Which paper grades do you most frequently purchase for fine-art reproduction jobs?	12. Which do you use most frequently for fine art reproduction?
Cur 1	Kodak NexPress (time donated)	N/A	N/A	Have used Mohawk Strathmore	Have used Mohawk Strathmore
Cur 2	HP Indigo	N/A	N/A	N/A	N/A
Cur 3	No Digital. Heidelberg for offset	Don't know. Have used Xerox iGen 4	N/A	Have used Xerox Digital Color Elite Silk; Burgo Chorus Art	N/A
Pub 1	Prints Ink-Jet	Epson	N/A	Epson Ultra Smooth Fine Art; Hahnemühle Photorag	Epson Ultra Smooth Fine Art

Interviewee	6. Which digital presses do you use most often?	9. What are the paper brands (manufacturers) you use most frequently?	10. Which paper grades do you most frequently purchase for non-fine-art reproduction jobs?	11. Which paper grades do you most frequently purchase for fine-art reproduction jobs?	12. Which do you use most frequently for fine art reproduction?
Pub 2	No digital . Heidelberg used most often	Gardicartierie, Scheufeulen, Gold East, Fedrigoni	N/A	Gardapat, Gold East Matte Art, Luxosant Offset, Phoenix Motion. Mori Silk	N/A
Pub 3	HP Indigo, Xerox Nuvera	HP packaged brands	GPA Matte	GPA Matte, HP matte	N/A
Pub 4	Kodak NexPress	Mohawk (by donations)	N/A	European coated bright white sheets	Printer makes choices most often
Pub 5	HP Indigo, Heidelberg for offset	Sappi (digital), Zantur, Gardipat (offset)	N/A	Sappi McCoy, Zantur	Sappi McCoy (digital) Zantur (offset)
Pub 6	HP Indigo	Mohawk	N/A	Mohawk Superfine, Cougar Opaque	Mohawk Superfine
Pr 1	No digital. Heidelberg and Komori offset	Sappi, Mohawk, Rolland, Burgo	Sappi Flow, Rolland PC100	Burgo Chorus Art, Sappi McCoy	Burgo Chorus Art
Pr 2	HP Indigo	N/A	100 lb text, premium sheet	100 lb text, premium sheet	100 lb text, premium sheet

Questions 7 and 8 were originally designed to aid in the selection of measurable predictors for the psychophysical experiment analysis, but these questions failed to provide useful data. There are several explanations for why these two questions failed to provided the expected results. First, the author assumed that the interviewees would be familiar with the different measurement terms. This was not the case. Although some did understand the terms, they were often understood by different names. Most distinctly, roughness was commonly referred to as 'roughness,' 'smoothness,' 'tooth,' and 'hand.' Gloss was referred to as 'gloss,' 'shine,' and 'luster.' Secondly, people selecting paper rarely focused on any specific attribute. Color is not evaluated separately from whiteness, brightness, or tint. Caliper is not evaluated separately from basis weight, opacity, or stiffness, and paper gloss is not evaluated separately from print gloss. Paper is most often selected based upon its integrated appearance while accounting for the type of book being produced and the cost. Finally, and perhaps most importantly, paper selection is a perceptual problem. Those factors selected in Q7 and Q8 are used as quality control metrics, not as perceptual predictors. Furthermore, they are common terms only to the paper and printing industries. It was an oversight to expect artists, curators, and publishers to, first, understand the terms, and, second, to rate the terms independently in decision making. Nevertheless, the results of the survey are shown in Figure 24.

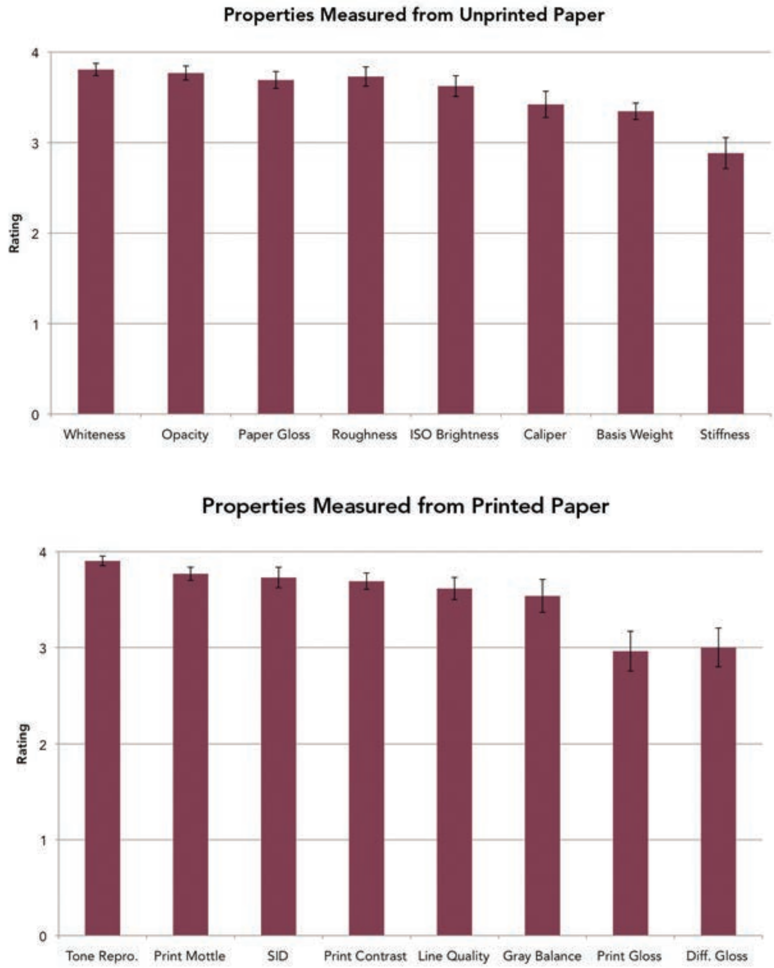


Figure 24. Collected results from survey of the importance of various paper and print quality metrics

Many interviewees defaulted to providing high ratings for all properties while some were more scrupulous in their ratings. The result was minimal statistical difference between both the unprinted and printed properties, although the highest rated properties were rated significantly higher than the lowest rated properties in both groups.

Open Discussion Analysis

Apart from answers specific to the prepared questions, the interviews resulted in many informative and revealing conversations about the nature of the printing, paper, and museum businesses. The major themes of those conversations are discussed in the following sections. The major themes were inspired by the interview questions and included discussions about finance, paper selection, relationships between people involved in fine art reproduction, and fine art reproduction workflows. The interviews provided a wealth of information about the nature of the fine art reproduction process, the relationships between people involved in the production of a book, and the problems and potential opportunities for the use of print-on-demand for fine art reproduction. The interviewees are not referenced by their name in this section, but are referred to using the coding system explained in Table 2. While the philosophies, methods, and business plans used to produce books of reproduced fine art are variable throughout the industry, the overarching goal remains the same, "… to try and find a balance and compromise between the reader's experience and enjoyment of the book, and, at the same time, reproducing the original images as closely and faithfully as possible" [Pub 6]. Of course, much is limited by the finances of the institution producing the book. The following section discusses finances from the interviewee's differing perspectives.

The Financial Influence on Paper and Fine Art Reproduction

Many museums are non-profit organizations. Their budget is allocated to maintaining the institution, while no money is allocated to the production of books. These museums rely solely on grants, such as those from the National Endowment for the Arts, or from private donors [CUR 1, CUR 3]. The amount of funding is variable. In some cases, there is much available or procured income for book production. This gives the institutions leeway in choice of paper, printing, and publication. However, and much more often, institutions are restrained by limited funding and must make decisions that will enable production of the highest quality book at the lowest possible price.

In some cases, museums will produce books simply because they would like to make available to the public a large collection of work not always on display. In one such case, CUR 3 oversaw the production of a book documenting an aspect of her museum's collection. This book was funded completely by government grants and private donations and required nearly four years to complete. This book was unique in that the publisher, in this case CUR 3, printed the book locally, rather than outsource to a foreign printer, who would have provided cheaper service with similar quality. By working locally, CUR 3 and her graphic designer were able to work hand-in-hand with the printer to ensure a quality product. Yet, cost remained an essential factor throughout the process, especially for paper selection. According to PR 1, "Cost is a huge factor for customers, especially when you're doing fine art books. Fine art books tend to be smaller runs. When you look at them per-piece they tend to be high. You have to get the quality and the value out of it. We want a great sheet but we want it to be cost-effective."

Cost influences paper selection for both the paper manufacturer and their clients. CUR 3 was looking for a high-quality sheet that could be delivered at low cost. At the same time, paper manufacturers are hungry for business. The local printer with whom CUR 3 was working had a working relationship with Burgo, the Italian paper manufacturer. It was apparent, from discussions with CUR 3, how eager Burgo was to sell them paper, in this case, one from their Chorus Art line. In the end, CUR 3 used the Burgo Chorus Art paper for her book, satisfied by both its price and the quality of the print.

Customarily, when a publisher decides to use a higher priced paper, that additional cost extends to the final price of the book. According to CUR 2, a publisher must ask "What's the quality of paper we can get and printing costs so we don't go over a certain amount and over-charge a person? ” For instance, some photography books, printed on heavy, matte paper, may cost $60 to $80. The same book printed on a cheaper, coated paper, would be less expensive. CUR 1 mentioned that her institution occasionally cut costs by using lighter paper than she would have liked. In such cases where cost is prohibitive, the printer can provide valuable insight. According to CUR 1, “They understand that we're dealing with a tight budget and they will often say that we've [the printer] got something that's almost like the one you like, but it's going to come in 25% cheaper,” and that she has “had projects where the printer said that they've got a lot left over from another job” and could provide it for a good price. Such compromise, between what the curator or publisher views as the ideal paper for the publication, and what is actually affordable or is more efficient for the production, is a common occurrence in fine art reproduction. In addition, according to PR 1, fine art books are printed in smaller runs, which means that they are predisposed to higher unit costs. Thus, it is ever more important to choose a paper that will balance the cost and quality of the book as a whole. Nevertheless, PR 1 stated, “It's usually pretty easy to do.” Printers have much experience working with clients printing on small budgets. In some cases, as described by PUB 4, the printer is told what papers to use and is asked only to quote the price. This may be the case if they are working with large publishing companies. In most cases, the printer's insight is well regarded. It is their recommendations and expertise that helps to ease the burden on curators and publishers to maximize quality and minimize cost.

Sometimes, though, paper selection decisions are made from ranks above those designing and printing the book. PUB 4 stated that she has “[t]rustees on the board [of her institution] that think if it's on glossy paper then it's expensive…When we were in the beginnings of the slump, I was told to stop using glossy paper because it looks expensive.” PUB 4 works mostly on small, in-house publications, such as exhibit brochures, fliers, and pamphlets. These small publications may see more day-to-day use than book publications, and thus, speak more first-handedly of the institution's philosophies.

There is precedent where an individual working at a museum or institution decides to produce a book as a personal project. In one such case, as told by CUR 3, the person in question, another curator, was passionate about bringing together poetry and works from the collection of the museum in which he worked. In this case, the book was published and funded independent of the institution to avoid burdening the institution with the responsibility of the publication, especially important considering the institution was not large and operated on a tight budget.

An additional financial incentive for paper selection is the use of ‘green' technologies. Publishers can advertise the use of recycled paper or Forest Stewardship Council (FSC) certified papers in their publications or the purchasing of wind credits by the printer to say the book was printed using wind power. According to CUR 3, “We are much more aware of the paper used for publications. It's kind of a point of pride now…how it was processed…how archival it is.” The up-front cost may be higher, but the marketability of using green processes pays off: hopefully, in sales for the publisher. The fashion of green processes is also illustrated in a case described by GD 1. She was working on a publication where the theme was “Going Green,” from an organization interested in advertising their transition to green processes. GD 1 incorporated a recycled paper into the design, and, in addition, the publication was dried using wind-power.

In many cases, books are published as marketing tools and to accompany exhibitions as merchandise. According to CUR 2, "Other institutions that you want to send shows to also want books. They need something for a book store. That's why artists today, of any stripe, or any curator, want to complete some kind of book, something that goes with that show because there are sales to be made in a book store." Herein lies the true financial motivation behind books of reproduced fine art. For a museum, books allow customers to bring an exhibition into their living rooms while providing the museum with supplementary funding for the exhibition, to help balance the cost of renting the exhibition and hopefully to provide some form of profit.

Up to this point, the cases discussed have been centered around offset lithography publications. An alternative is the use of digital printing technologies and POD workflows. In POD workflows, content is loaded into a printing queue, then printed (on digital presses), bound, and shipped with minimal human handling. The advantage of using POD workflows, such as those used to produce photo-books by companies such as Kodak, Blurb, Lulu, and HP, is that a single book can be produced one-off. While the unit cost is higher than for offset lithography, according to PUB 6, "Sometimes it makes more economic sense to pay a higher unit cost for a few hundred copies than it does to pay a smaller unit cost for a few thousand. The last thing you want to have happen is to end up with large numbers of unsold copies because…you still have to manage inventory" (author's ellipsis).

However, the business model for POD can vary. Papers used in digital presses are specifically designed for particular press technologies, such as those used by HP, Kodak, and Xerox. In the high output commercial POD industry, the printer has the most influence over paper selection, according to PR 2. Commercial POD workflows rely on efficiency. Digital presses contain trays for a small number of paper types for which the press is calibrated and controlled. One local POD printer produces thousand of books in a single day. Thus, to maximize efficiency, the printing company has selected a small number of papers from which customers can choose, thereby minimizing the number of times papers must be exchanged in a press. Those selected papers were chosen for use on a variety of jobs. For the printer to decide to introduce a new paper, according to PR 2, either large organizations known to bring in sufficient business will have to back the decision, or the process would have to be arranged such that all orders to be printed on that paper would be printed at a specific time, non-disruptive to general operations.

If POD workflows are designed appropriately, they have the potential to be a valuable resource for museums and artists because they minimize inventory. Ideally, a customer could order a book from a museum or a specific artist and have it printed at the time of purchase. According to PUB 6, "Some great books that could have been printed don't end up being printed because it becomes too expensive…Digital printing produces high-quality controllable color work…I'm confident that this is going to be the way to go." However, for this model to work with large commercial POD printers, new papers would have to be introduced into the workflow, unless the museum or artist agree to use those already available. The introduction of new papers into a commercial POD workflow "[c]osts time and energy, and there's inventory involved in warehousing it" [PR 2]. Alternatively, small POD printers, such as the RIT Cary Graphic Arts Press, have more leeway in paper choice because they publish fewer books and have a history of producing short run, high-quality, fine art books. The problem is best summarized in a statement by PR 1, who, incidentally, works more with fine art reproductions produced using ink-jet printers than digital presses, "[y]ou have to make sure you can get all of the factors in: reproduce the images, do it within budget, do it on time, and again if they do a reprint." The complete business model, including design, publishing, paper selection, and printing, is dependent upon the ability to print a book, with good acceptable quality, in time, and within budget.

Paper Selection

The influence of finances on paper selection was discussed in the previous section. While paper cost plays a large part in determining the final choice for paper, many other factors are incorporated into the paper selection process. Among the most important factors influencing paper selection for books of reproduced fine art is how the paper represents the art medium being reproduced and the integrity of the final image with respect to the original. In some cases, this is best achieved by selecting a paper that mimics the surface of the original work of art. In other cases, paper is chosen to produce the best reproduction of the digital image of the original work. GD 1 admitted that, while she was not solely responsible for selecting paper in her position, that "[t]he period that the image is from…has some sort of influence on the type of paper" (author's ellipsis) they used. CUR 2 discussed a case in which a book of photographs by 20th century photojournalist Bernie Boston were reproduced. The fact that most of his photographs were reproduced on newsprint was considered in the decision-making process, with an understanding that newsprint would not be used in the book. CUR 1 went so far as to say, "[w]hat we're trying to do in almost every case is to reproduce a photograph as accurately as possible. That means paying attention to the kind of paper the original was printed on, the surface of the photograph, the contrast, all of those things are important to us. I think it's the most important consideration for us." CUR 2 put it well in her comment on printed reproductions of photographs:

> It's the ground on which the image rests. I don't lie to myself, that this is ink on paper. This is not a photograph. It's a reproduction. What you want is that reproduction to be 'as close as' that thing itself…The reproduction needs to be clear, informational, it should not try to change the effective quality or the concept of the work. Anything that comes in contact with the reproduction cannot do it as well. It needs to be supportive of the original work. We're playing with abstractions. That's what's interesting about what we do. Trying to create a similar original.

Reproductions of art are not meant to be facsimiles of the original artwork. CUR 3 makes her paper decisions, both aesthetically and technically, by taking "into account the original medium in terms of the feel of the paper and in terms of printability and how the ink will lay down, in contrast to what's happening on the original artwork." Both offset and digital technologies have difficulty printing on materials similar to canvas, yet that is the material used for many paintings. Therefore, the best option is to choose a paper that will best render the photographic reproduction of the original artwork. However, while there may be few options for paper surface, there is freedom in color choice. PUB 3 stated that she prefers a very neutral paper when reproducing primarily dark paintings, and when reproducing photographs, the paper color may vary depending on the color of the original artwork. PUB 4 also discussed how, for example, a cebachrome photograph is best reproduced on high gloss paper.

Nevertheless, the common theme was that of compromise. Books of reproduced fine art often contain reproductions of several different art media. Therefore, it is necessary to choose a paper that will globally satisfy the reproductions included in the book. This means that paper is selected to best reproduce the images of the art work, not to best represent each and every art media. PUB 6 stated, "[w]e have to print it [a book] on a paper that lends itself to the best reproduction techniques…Obviously the aim of any publisher who's reproducing art is to make the best possible reproduction and the very summit of that would be to reproduce extremely faithfully all of the color values and tonal values of the original…We would like to be as color and tonally accurate as we can…," but he went on to say that, without choosing a paper that is very white and bright, readability would be negatively impacted. CUR 3 agrees that "[w]hat would feel good to the reader" is an important factor, and that specifically, "[t]here shouldn't be very much texture because that definitely would interfere with how you see the image and the text." The thickness of the paper was also mentioned as an important factor. Thus, among the many factors that influence paper selection, the one chosen may be the one with the best ink holdout, highest printed densities, largest contrast, most appropriate color for the collection, and/or the one providing the best readability.

Despite the considerations described above, paper selection decision is still largely a business decision. In some cases, the people making the final decision do not always have print quality as their number one objective. PUB 1 (who works primarily with inkjet reproductions) discussed a case where she made reproductions for a Vermeer show. Her boss anticipated a large quantity of sales and, thus, wanted to print on a thinner, less expensive paper. Vermeer paintings are low key and, thus, require a lot of ink. She noticed that there was much mottling in the shadow areas of the images, most likely due to the use of a thinner sheet with poor formation, a problem that may have been solved by using a higher quality paper. Nevertheless, the images sold well, a further illustration of the disparity between an artist's eye and that of a customer. In this case, the manager may have known about the problems caused by using a less expensive sheet, but made a decision based upon knowledge of the customer base. In most cases, said PR 1, "[i]f you've got a lot of ink coverage you cannot run a sheet that does not have good opacity or a sheet that's not very heavy." The necessary quality of a sheet really depends on the specific job at hand.

Factors relating to paper permanence and archivability may also be considered in paper selection. CUR 3 stated, "[t]he other parameter that we're working with is what libraries now want, the standards that libraries are…demanding…It seems to me that libraries are now asking for certain levels of paper quality so that the paper will last for archival purposes." At the time of the interview with CUR 3, this author was not familiar with archiving standards and did not further question CUR 3 on the subject. However, after conversations with Douglas Nishimura of the Image Permanence Institute at the Rochester Institute of Technology, it became apparent that this was a hot-button issue within the paper and library industries. In 1984, the American National Standards Institute (ANSI) produced the standard ANSI Z39.48, called "Permanence of Paper for Printed Library Materials" (D. Nishimura, personal communication, April, 2011; McCrady, 1998). The standard, later revised in 1992 and adopted by the National Information Standards Organization (NISO) in 1995, included a wide variety of books in its scope. The standard's purpose was to establish "[t]he criteria for permanence of uncoated paper and it meant to ensure sufficient longevity of the paper so that it should last several hundred years under normal conditions of library circulation and storage without significant deterioration…" (ANSI/NISO Z39.48, 1992, p. 1). It included specifications for minimum pH, cross-direction folding endurance, tear resistance, minimum alkaline reserve, and lignin content in paper stock (fold endurance was removed in the 1992 revision) (McCrady, 1998). The standard was constructed as such because libraries complained that lignin caused the paper to both yellow and become brittle. In a 1996 meeting the paper industry contested the claim that lignin cause both yellowing and brittling. They admitted that yellowing was caused by lignin content, but claimed that the brittle paper was caused by the acid paper-making methods and the alum rosin sizing (D. Nishimura, personal communication, April, 2011. The paper industry, correct in their assertion of what caused paper to become brittle, was motivated to defend lignin because removing it in the pulping process reduced the paper yield per tree. Nevertheless, ANSI Z39.48-1992 is still referred to in the cultural heritage field, but whether it is a sufficient standard is still up for debate.

The archival nature of paper stocks not only influences the decision of paper purchasers, but is also an important tool used in paper manufacturing. PM 1, in discussing one of his company's most popular papers, cited archival quality as important factor in its popularity. They took into account many different archivability standards when creating the paper, be they from ANSI, ISO or the Library of Congress. Nevertheless, PM 1 admitted that paper selection is still heavily dependent on the nature of the art being reproduced. The same paper grade, popular for being archival, is also popular because of its slight texture. PM 1 stated about producers of books, "[s]ometimes they want something that has a little more texture, and tooth, and hand to it if you will." In certain cases, the artist might require a very bright, blue-white, glossy paper, for creating images with great contrast and saturation, such as graphic art or pop-art, although, "[a]rchival properties and a nice toothy surface is what historically has served the fine art reproduction market well" [PM 1].

However, after all is said and done, technology remains the limiting factor in paper selection, especially if the book will be printed digitally. Toner particles are prone to be unevenly distributed on rough surfaces or to fuse poorly to surfaces without the proper treatment. This is where the expertise of the printer becomes important. They are aware of what papers will work on their equipment and can make recommendations, not only for cost but for what papers will have the best printability. “From a manufacturing perspective,” according to PR 2, “you can make a book out of just about anything. The equipment’s fairly steady. Once you actually have the physical book block that’s printed, the manufacture of the book is not much of an issue.”

Relationships Between People and Production Workflow

Paper selection is an important component in fine art reproduction systems. Equally as important are the relationships between the people working within the system. There are many types of people involved in the production of fine art reproduction books: printers, publishers, curators, designers, and managers. Often, one person may play multiple roles. This is counter to the discussion from the introduction of this thesis, where printers, publishers, artists and users were hypothesized to be those most involved in the fine art reproduction process (see Figure 1). All interviewees agreed that users have zero influence on paper selection, unless curators are classified as users when conducting scholarly work. In most cases, living artists also have little to no influence on paper selection unless they are personally involved with the project, either as a publisher or a participant invited by the publisher into the process.

Many of the interviewees discussed the important role played by graphic designers in the paper selection process. PUB 5, after hearing the initial list of people discussed in the introduction of this thesis, was adamant that designers be included. CUR 3 discussed the role their designer played in the creation of their book, specifically citing her long-standing relationship with local printers. The designer is often the visionary behind books of reproduced fine art. Therefore, they work with the printers to ensure color accuracy and to make sure that the product appears how they had envisioned. PUB 6 stated that, “[t]ypically in the publishing process it’s the designer that recommends the paper. We might use a freelance designer, so we would trust the judgement of the designer.” The paper market is quite vast. Publishers are most concerned with the production of the book itself and may not have the time to devote to paper selection. It is easiest for them to turn to designers or printers, who deal in paper on a day-to-day basis. Yet, printing technology changes so rapidly that it can become difficult for designers and publishers to adapt their workflows. CUR 3’s designer was trained in an era when only traditional printing was used. The market for digital printing has increased rapidly since its beginning, and thus, paper manufacturers have introduced many new products into the market where each paper is specifically treated for a particular press. Therefore, it may be difficult for older designers to predict “[t]he way things will look when they’re digitally printed,” which may “change completely her choice of papers or papers that she’s comfortable with” [CUR 3]. Green products, discussed earlier, are included among those changing technologies.

The assistance printers provide for clients printing on a small budget was discussed earlier. However, printers play a valuable role in paper selection in a more general sense. PUB 3 discussed the role of printers when using digital printing technology, “Because digital printing is so new, and because [printers] have access to these papers and the most knowledge about these papers, we depend on them entirely for their suggestions." GD 1 said in support, "[t]he printer plays a critical role because they check on the availability of the paper for us or they might recommend something different…," or "[t]hey might make a suggestion based upon the price point." Publishers often may have an idea of what they would like the paper to be, but it is the designer and printer who understand the true functionality of the paper and the feasibility of using one over another. For printers, it is in their best interest to have a broad knowledge of papers and the ability to accommodate many budgets due to the large number of printers looking for business. As with any contractor, PUB 5 states, "[w]henever we have a job we send them out to our core printers, then review the estimates and schedules and pick the best price and best schedule." Considering that PUB 5 works at a large institution (the Smithsonian), it is likely that his relationship with printers is less personal than that of CUR 3, who comes from a local Rochester museum. Options are most limited when working with large publishing companies such as Taschen, as described by CUR 2, "[a] lot of publishers have set templates and the paper's set too. It's about cost effectiveness… you do things like that that are templated-out…You put the images in and you write to what's left on that page. They're going to go with a paper that's more cost-effective, but thick enough so that you don't read one image on the other side." Large publishing companies have runs where thousands of books are produced. Anything upsetting the carefully streamlined workflow is rarely acceptable. This sentiment is similar to PR 2's comments on the difficulty of introducing a new paper into a commercial POD workflow. Unless it brings in enough income, it only served to upset the workflow.

PR 2 discussed two different models in POD: consumer-driven content, where "[c]onsumers have their own content and they want to do something with it in a hard copy output," and organization-driven content, where organizations have "[i]mages people want to see, zoos for example, and they are just sitting on that content." Such is the potential market for POD and digital printing in general. Still, digital printing is under-utilized by the cultural heritage community. Both CUR 1 and PUB 3 admitted to using digital printing mostly for small publications, such as research journals, brochures, post cards and announcements, while CUR 1 stated that they were steadily using digital printing for book publications. Of those available digital printing technologies, PR 2, PUB 3, and PUB 6 stated they primarily used the HP Indigo, although they did not mention the particular models. The local commercial POD company also used digital presses from, Kodak, Xerox, Xeikon, and Canon, each with different technologies and serving different production purposes. There is much competition between digital press manufacturers as PR 2 describes, "[w]hat's interesting is the [Xerox] iGen has a much finer toner particle than the Indigo, so if you look at the marketing material coming out of Xerox…their images are hypersharp and the detail on them is exquisite, but they design their images…with the purpose of being run on the iGen versus that same image being printed on the Indigo." which will, " [l]ook fuzzy and blurry because the Indigo is not able to hit those fine tonal reproduction qualities." However, as he went on to question, is that amount of detail really relevant to the bulk of digitally printed content? All digital presses reproduce sharp text and most consumer images are dominated by low frequency information.

Nevertheless, paper manufacturers must always be inventing new products and improving on their current products to account for changes in market demand. PM 1 commented on this process specifically for fine art papers. It was previously discussed how papers are manufactured using different processes for different digital presses. PM 1's company surveys both digital press manufacturers and paper buyers to understand how they can improve current products and what properties are important for inclusion in new products. Equipment manufacturers may go to the company and say "[w]e're looking at a new technology. Can you help us design a product that will optimize it? " [PM1]. In the research and development department, PM 1 develops new products by answering the questions, "what does the market need, what do we have the ability to do to meet that need, and how would we go about doing it?" It is the constant interplay between paper manufacturers, press manufacturers, and paper users that keeps the paper and print industries constantly improving to meeting customer demands, and it is why digital presses are even being considered for fine art reproduction today.

Psychophysics Analysis

The psychophysical experiment dataset consisted of many components. Paper and print quality was analyzed for each paper and print combination using a variety of standard tests. The complete set of measured data was considered as the set of factors by which the various psychophysical responses were analyzed. The psychophysical responses included Color Rendering Quality (CQ), Surface Appearance Quality (SQ), Image Quality (IQ), and Customer Preference. As was previously discussed, the psychophysical experiment design was predicated on the principles of the Image Quality Circle. Therefore, the psychophysical data analysis was also constrained to the principles of the Image Quality Circle. The physical measurements defined the PIPs, and CQ and SQ data defined the Customer Perceptions and the image quality rankings defined the Customer Image Quality Ratings. Visual algorithms were constructed to estimate Customer Perceptions from PIPs and image quality models were constructed to estimate Customer Image Quality Rating from Customer Perceptions. Most of the analysis strictly adhered to the latter order of operations; however, for the sake of thoroughness, an abbreviated Image Quality Circle model was also tested whereby Customer Image Quality Rating was estimated from Physical Image Parameters.

The first experiment discussed in this section is the Coolness Estimation Experiment. The designed psychophysical experiment contained four predictors: coolness, PST, roughness, and gloss. Of these four predictors, coolness was the only factor that could not be physically measured using an already existing device. Thus, before discussing the psychophysical experiment analysis, it is necessary to define the coolness metric, and thus complete the set of designed experiment factors. The three Image Quality Circle models will be discussed following the discussion of the Coolness Estimation Experiment. The Image Quality Circle models were divided into three sections: Visual Algorithms, Image Quality Models, and Direct to Customer Image Quality Rating Models. Within each section, models based completely on the designed experiment—meaning they include only coolness, PST, roughness, and gloss as parameters—and expanded models— meaning they contain additional measured factors—will be discussed. For the purposes of brevity, the various models are given abbreviated names, described in Table 14.

Table 14. Model names and abbreviations

Model Name	Model Code
Visual Algorithms	
Designed Color Quality Model	DCQ
Designed Surface Quality Model	DSQ
Expanded Color Quality Model	ECQ
Expanded Surface Quality Model	ESQ
Image Quality Models	
Designed Image Quality Model	DIQ
Expanded Image Quality Model	EIQ
Direct to Customer Image Quality Rating Models	
Designed Direct to Customer Image Quality Rating Model	DD2IQ
Expanded Direct to Customer Image Quality Rating Model	ED2IQ

Paper Measurement Analysis

Table 15 contains paper quality measurements of roughness, caliper, basis weight, opacity, ISO Brightness, CIE Whiteness, and CIE Tint for the twelve papers included in the psychophysical experiment. The table illustrates the large amount of variation between the twelve papers. For example, Paper D is very rough, with a Parker Print-Surf of 7.73 and a Sheffield Smoothness of 333.00, very thick, with a caliper of 7.8 mils, and low gloss, with a value of 2.23. Paper F, on the other hand, is very smooth, with a Parker Print-Surf of 1.03 and a Sheffield Smoothness of 27.00, is relatively thin, with a caliper of 4.48 mils, high gloss, with a value of 35.90. Both Papers D and F are 100 lb. text papers and are have similar opacity and brightness values.

Table 15. Paper quality data for the 12 sample papers[17]

Paper	Parker Rough.	Sheff. Rough.	Caliper	Basis Weight	Opacity	Brightness	Whiteness	Tint	60° Gloss
A	5.03	86.6	5.04	81.36	92.7	**98.71**	**126.49**	0.17	2.97
B	4.37	65.5	6.08	100.61	94.38	98.69	125.86	0.15	3.07
C	7.22	253.5	5.98	81.45	92.46	94.65	121.6	0.33	2.2
D	**7.73**	**333**	**7.8**	99.98	96.3	94.86	123.19	-0.04	2.23
E	1.1	**25.4**	**3.46**	81.34	94.64	91.53	114.29	-0.52	**37.83**
F	**1.03**	27	4.48	103.87	95.94	92.35	115.89	0.07	35.9
G	4.54	68	4.82	83.46	**91.1**	84.37	79.05	-2.93	3.93
H	4.33	60	5.94	101.94	93.96	**84.19**	**76.99**	-2.99	3.93
I	7.63	326	6.38	82.93	94.78	79.36	63.71	**-4.74**	**2.2**
J	7.18	266	7.7	102.6	93.82	89.79	77.83	-2.14	2.4
K	1.71	27.5	4.12	83.04	94.14	92.99	110.96	**0.9**	11.57
L	1.62	28.6	4.82	103.09	**96.48**	92.35	109.1	0.8	19.57

It was previously discussed how basis weight was used as a confounding factor with opacity considering that opacity between, for example, 80lb text stocks, differed very little. Notice that, for all papers except Papers I and J, papers with lower basis weights also had lower opacities. However, it was also discussed how opacity is a poor indicator of bleed-through, a factor of higher importance than opacity when evaluating the ability to see text or images on a duplex printed sheet. Print Show Through (PST), the metric developed at International Paper (J. Kohler, personal communication, January, 2011), enabled the evaluation of text and image visibility on duplex printed sheets. The difference between PST and opacity is illustrated in Figure 25, where PST is plotted against ISO Opacity (both normalized by their mean value to aid in visual comparisons) for the twelve papers with prints made using the HP and Kodak presses. While there is a strong linear relationship between PST and opacity for papers with very low and very high opacity, the relationship is blurred for papers with mid-range opacity.

17 - Values shown in bold represent the highest and lowest values across the set of papers.

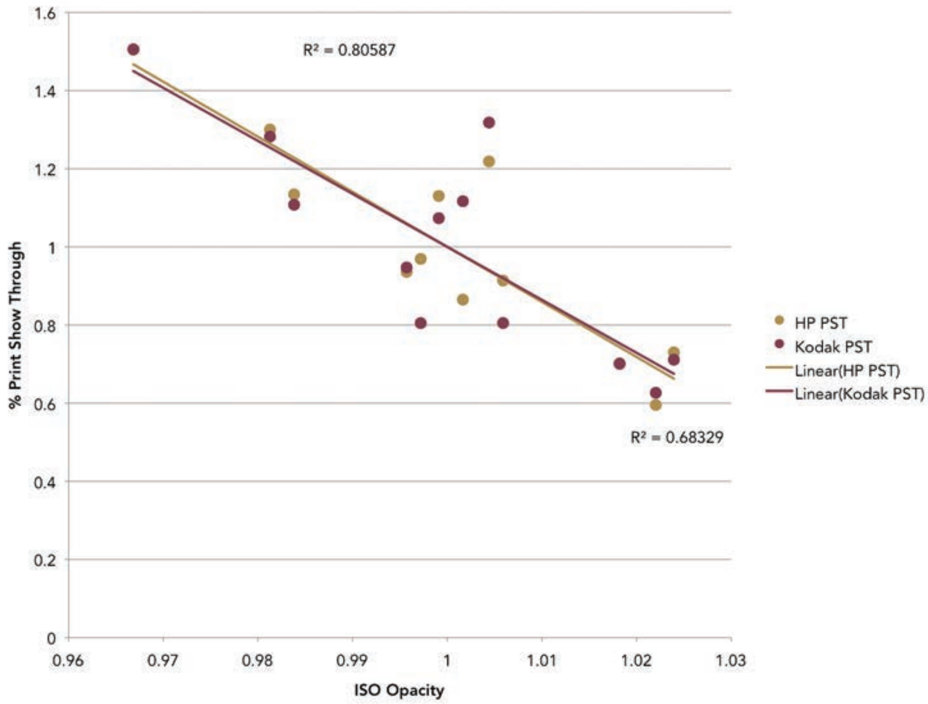


Figure 25. PST plotted against opacity with both metrics normalized by their means

Figure 26 shows PST for the twelve papers with images printed using the Kodak NexPress S3000 and HP Indigo 7000, the two presses included in the experiment. Interestingly, as with opacity, Paper I has a higher PST than Paper J. While Paper J is a 100lb sheet, it has a lower roughness and a higher caliper than Paper I, suggesting that, while it may have had greater bulk, it contained less filler or other components added by the manufacturer to boost opacity. Several different papers were chosen to increase the variety of papers used in this study; however, many papers, such as Papers A and B and Papers K and L, were the same stock but differing in basis weight. This ensured consistency between measurements of the two papers and that differences between them would be most likely attributed to differences in basis weight. This comparison could not be made between Papers I and J because they were different stocks, although from the same manufacturer. Herein lies one of the difficulties of designing an experiment for the study of papers while using only commercially available stocks.

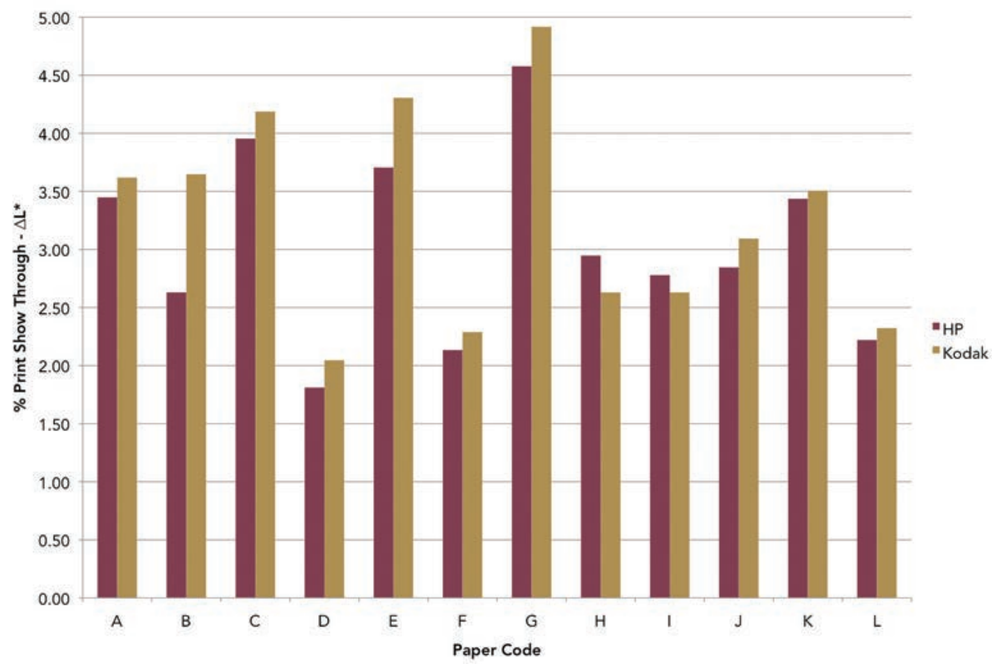


Figure 26. PST for the 24 prints

The paper and printing industries commonly use two methods, Parker Print Surf and Sheffield Smoothness, for the measurement of roughness (or smoothness, as some prefer). Sheffield Smoothness is most commonly used to measure uncoated papers, while Parker Print Surf is commonly used to measure coated papers. However, while both coated and uncoated papers were included in the study, it was necessary to choose a single metric. Parker Print Surf values are commonly observed between 1.0 and 4.0, where smoother sheets result in smaller values. Sheffield Smoothness values can vary between 0, for the absolute smoothest sheet, to 500, the roughest possible reading for the system. However, the Parker Print Surf device digitally displays the roughness value, while Sheffield Smoothness is judged by the position of a plastic pin floating in a glass tube, the height of which is controlled by air pressure within the tube. The Sheffield Smoothness device is very difficult to calibrate, does not have the resolution of the Parker device, and tends to have higher variability for rougher samples. Plots of the roughness standard error and mean roughness for the 12 experiment papers are shown for Parker Print Surf in Figure 27a and for Sheffield Smoothness in Figure 27b. It is evident from these plots that Parker Print Surf is more stable across a range of coated and uncoated papers than Sheffield Smoothness. Therefore, Parker Print Surf was used as the metric for describing roughness in this experiment.

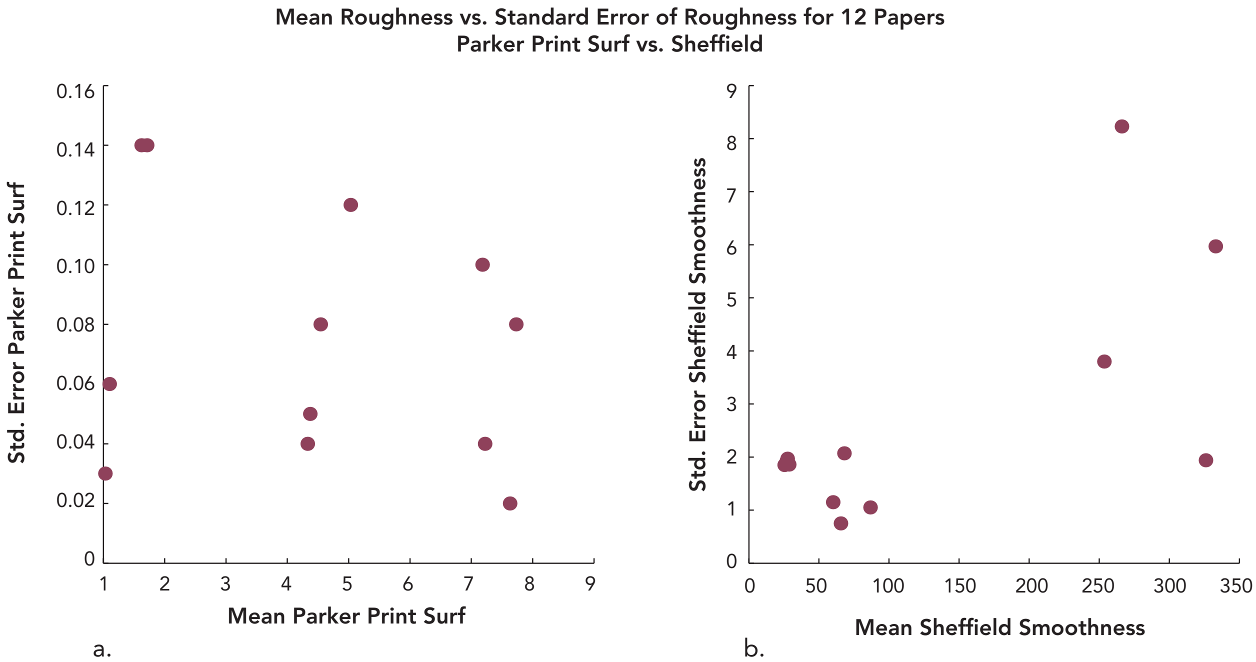


Figure 27. Plots of (a) Parker Print Surf standard error versus mean Parker Print Surf and (b) Sheffield Smoothness standard error versus mean Sheffield Smoothness for the 12 sample papers

Furthermore, there is a highly predictable relationship between Parker Print Surf and Sheffield Smoothness, as shown in Figure 28. Sheffield Smoothness can be described by Parker Print Surf through the use of a power function (the equation is shown in the figure) with an adjusted R^2 value of 0.9984.

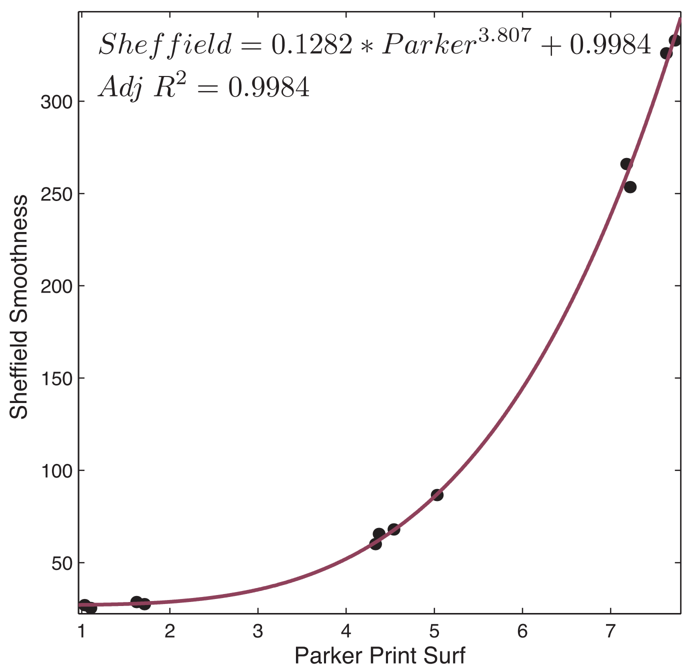


Figure 28. Plot of Parker Print Surf versus Sheffield Smoothness with the exponential fit relating Sheffield Smoothness to Parker Print Surf

The results of the Coolness Estimation experiment are described in the following section. This experiment developed the coolness metric, the final predictor from the four used in the designed experiment.

Coolness Estimation Experiment Analysis

The Coolness Experiment was designed because current metrics used by industry for describing paper color do not adequately describe the sensation of a paper being 'cool' or 'warm' colored. Figure 29 shows the Normalized Scale Position (NSP) of the 12 experiment papers plotted against three metrics commonly used to describe paper color: ISO Brightness, CIE Whiteness, CIE Tint. In addition, a linear regression between CIE Whiteness and CIE Tint was constructed and plotted against NSP because they are commonly used in conjunction to describe paper color.

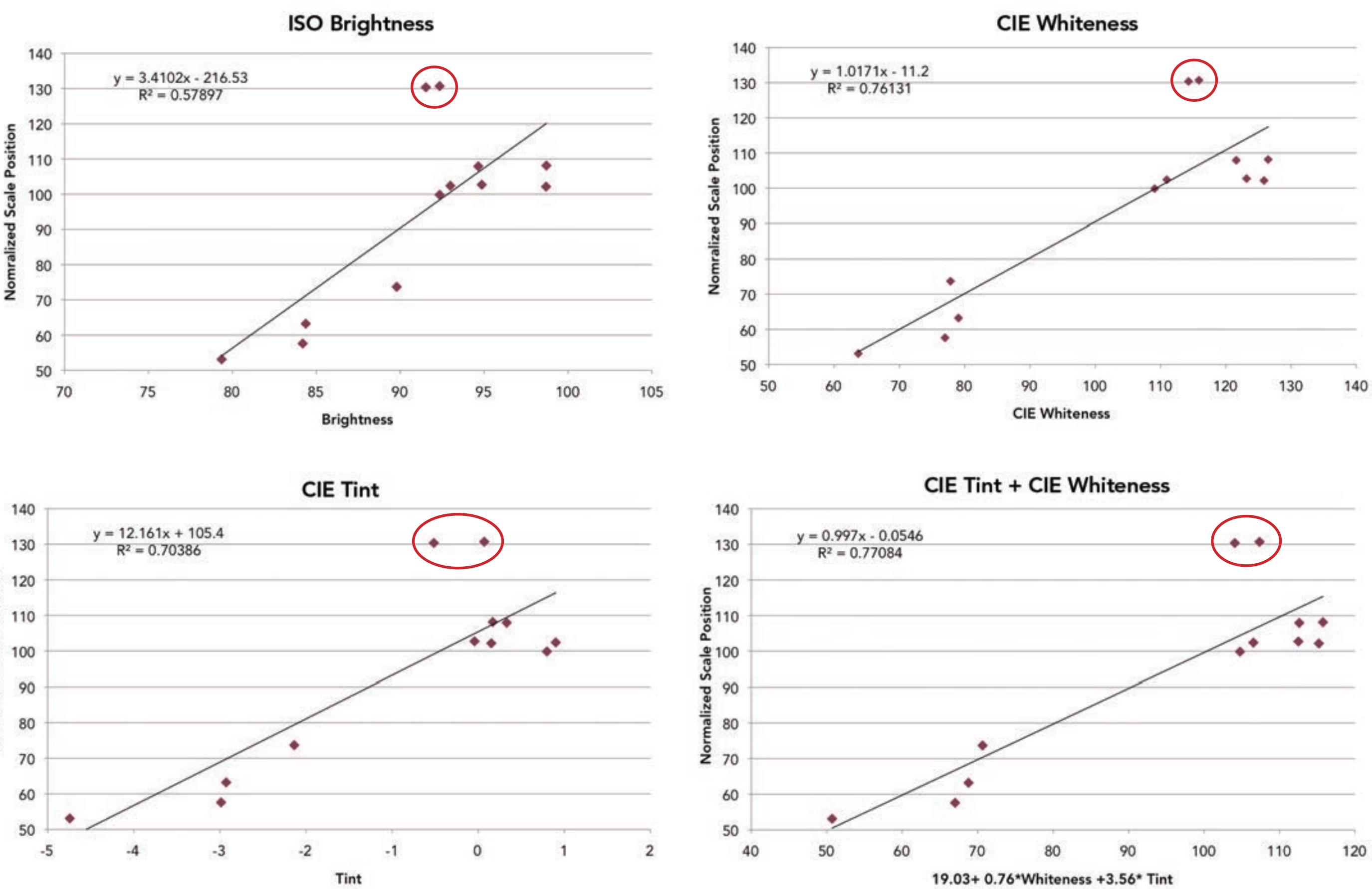


Figure 29. NSP versus ISO Brightness, CIE Whiteness, CIE Tint and a linear combination of CIE Whiteness and Tint for the 20 coolness estimation experiment samples

Beside the fact that ISO Brightness, CIE Whiteness, and CIE Tint were not designed to quantify the color of papers containing optical brightening agents nor to quantify color under non-standard conditions, the four plots in Figure 29 all display the same error. The red circle in each plot encloses two points judged to have very high coolness values. Each of the four tested metrics greatly under-predicts the perceived coolness of the two samples. Despite the two under-predicted points, the linear fits between the four metrics and NSP were still less than optimal.

The decision was made to develop a new metric for coolness based upon CIELAB. Coolness is conventionally thought of as the differentiation between hues of cool colors—namely blues, cyans, magentas, and sometimes greens—and warm colors—namely reds, oranges, and yellows. However, white paper is, by definition, not of high chroma. Paper manufacturers add dyes and optical brightening agents to paper pulp to produce warm or cool shades of white. The variability among paper colors is relatively consistent. The CIELAB a^* and b^* values (2° standard observer under illuminant D50) for the 20 papers included in the Coolness Estimation Experiment are shown in Figure 30. The papers are seen to vary consistently along an a^*b^* vector, primarily along the b^* axis with a slight change along the a^* axis.

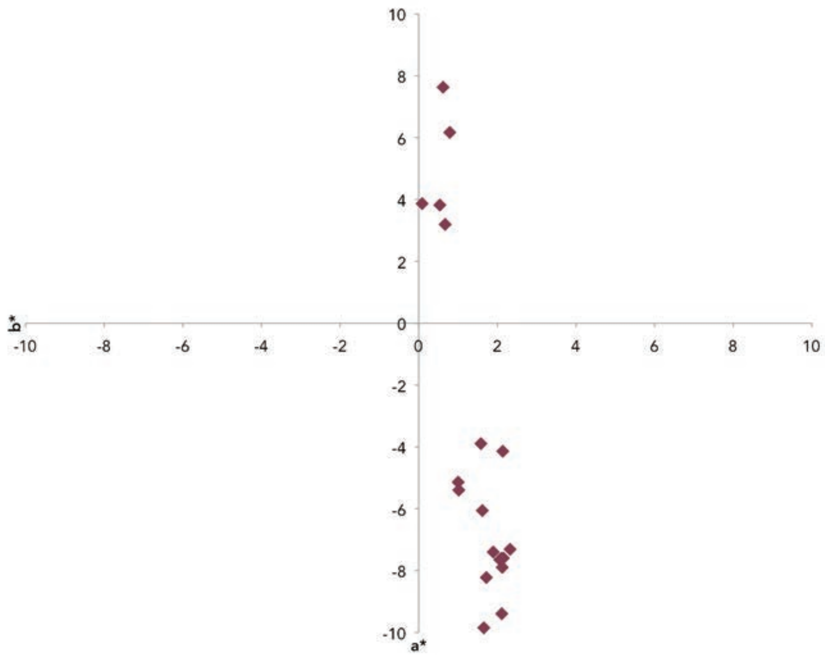


Figure 30. CIELAB values for the 20 paper samples used in the coolness estimation experiment

Cool paper shades are created by the addition of blue shading dyes and optical brightening agents to the paper pulp. Increased concentrations of optical brightening agents result in decreased b^* and increased a^* values. Changes along this vector may be adjusted, amplified or controlled by the shading dyes. Conversely, papers with warm shades are created completely by the addition of dyes to the paper pulp. Thus, because paper color varies more by chroma than by hue, it was hypothesized that the perception of coolness would be best described by a metric based upon C^*. However, because paper color was known to change most significantly along the b^* axis, an additional constant was added to C^* such that the sign of C^* depended upon b^*. Several variations on the C^* equation were tested against NSP for the prediction of coolness. However, the simplest variation proved the most pragmatic. The metric used to describe coolness in this experiment is expressed by Equation 14. The metric is called *ChromaV1*, becuse it was the first tested variation of chroma and to leave open the possibility for further variations if necessary.

$$ChromaV1 = \frac{b^*}{|b^*|}\sqrt{a^* + b^*} \tag{14}$$

At first, *ChromaV1* was tested using measurements made with the UV component included, with the 2° standard observer and under illuminant D50. The 12 experiment papers were used to train the equation (although little training was actually needed) and the 8 supplemental papers from the Coolness Estimation Experiment were used to test the equation. Figure 31a shows a plot of NSP versus *ChromaV1*. Surprisingly, *ChromaV1* appeared to under-predict the coolness of the same two samples rated with the highest coolness (shown by the red circle in Figure 31a), as were under-predicted by the four metrics in Figure 29. Additionally, the test data suggests some nonlinearity between NSP and *ChromaV1* (see Figure 31b). This evidence means that *ChromaV1* is a poor metric for describing coolness. However, *ChromaV1* was recalculated using CIELAB measurements with the UV component excluded (measured using a Gretag Color-Eye 7000). This slight adjustment to the measurement procedure appeared to solve the latter problems completely. Figure 31c shows a plot of NSP versus *ChromaV1* with the UV component excluded. The two points, shown in the green circle, of which the NSP was previously under-predicted, are now in alignment with the remaining data with an R^2 of 0.97. In addition, the test data are also linearly aligned (see Figure 31d). The precision of the fit between NSP and *ChromaV1* for 100, out of 125,970 possible combinations of 12 papers from the 20 available, was also tested to verify the legitimacy of *ChromaV1* as a metric. The mean R^2 for a linear fit was 0.96. Thus, *ChromaV1* with the UV component excluded was used as the metric for describing coolness in this experiment.

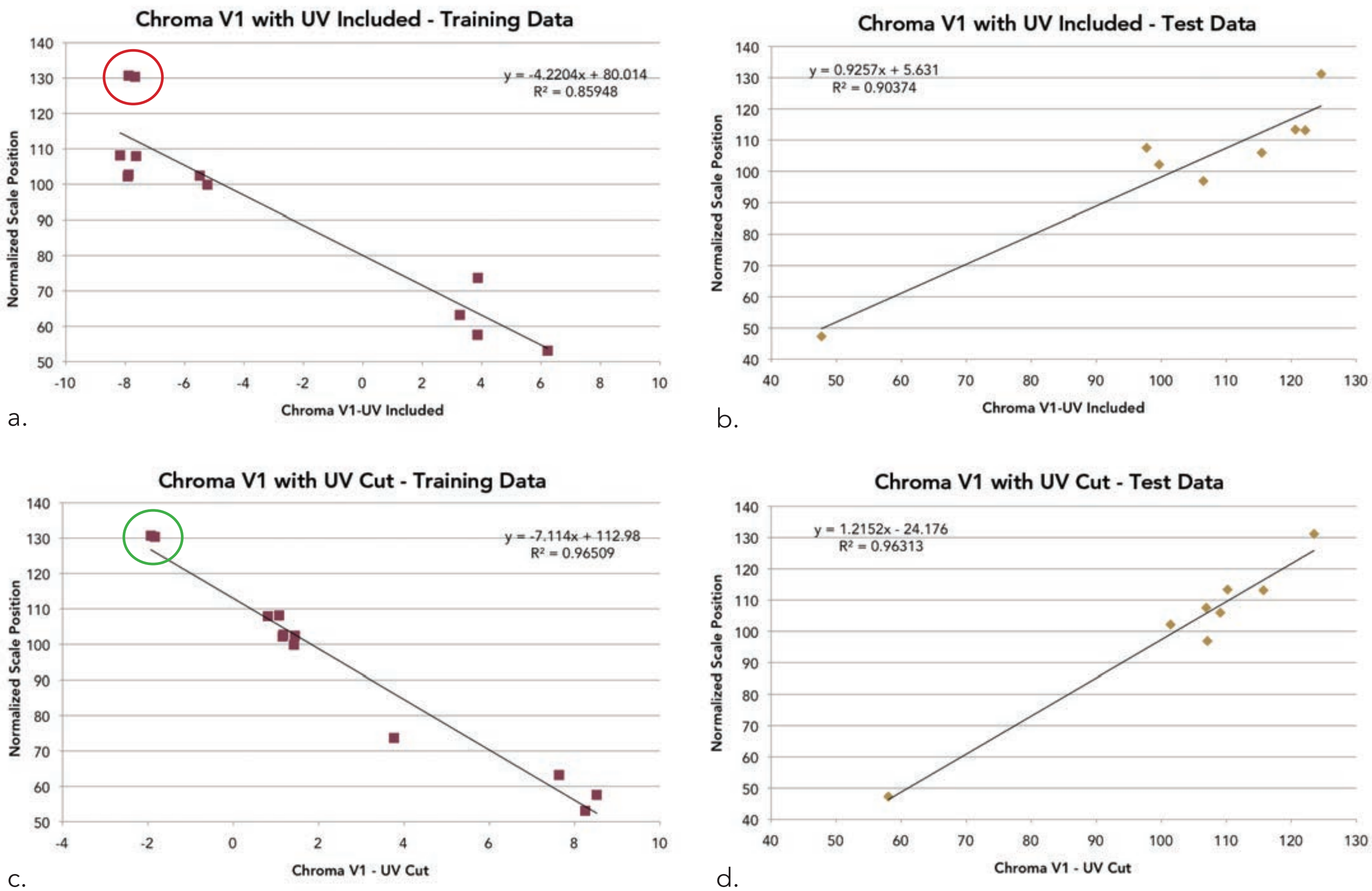


Figure 31. NSP versus Chroma V1 for: (a) UV included with the training data, (b) UV included with the test data, (c) UV cut with the training data, and (d) UV cut with the test data

The perceptual consequences of the Coolness Estimation Experiment results should not go unmentioned. No prior research has come to the attention of this author regarding experiments examining the perceptual scaling of white paper. Assumptions are often made about the perception of optical brightening agents, that they make the paper appear 'brighter.' This is underscored by the presumption that bluer objects appear brighter. Thus, it is curious why the perception of coolness, a term used so often by artists, printers, and designers to describe paper color and a perception heavily influenced by the optical properties of paper, should be better described by UV excluded measurements than UV included measurements. The twenty papers were all viewed under simulated D50 fluorescent lighting, a source with enough UV energy to cause considerable excitation of the papers' optical brightening agents (for those that have them). Yet, the metric best describing the perceptual results was based upon a measurement system in which the optical brightening agent excitation was neutralized. While this phenomenon is not the focus of the research, it is of interest to the author and will be studied in future work.

At this point, the four sample selection parameters included in the designed analysis model and the additional predictors included in the expanded analysis model have been defined. The following sections will describe the statistical analysis of the potential models befitting the Visual Algorithm and Image Quality Model components of the Image Quality Circle.

Response Definition and Physical Image Parameter Selection

Before discussing the statistical derivation of Image Quality Circle components, it is necessary to describe the method with which response values were derived and with which PIPs were selected. The derivation of the responses attributed to Customer Perceptions and to Customer Image Quality Ratings are discussed below.

Response Definition

Probability of Selection

All data from the psychophysics experiment was collected using the rank order method. The first step to defining customer responses from rank order data was to determine the probability a particular book was ranked in a given position. An assumption was made that each observer who participated in the experiment was selected from a random population because they volunteered for the study and were not specifically singled out. While this was not random in the strictest definition of the word, it is the most practical random selection procedure for the university environment. Therefore, it was assumed that the randomness of observers extended to randomness of book selection. For example, the probability that Book A would be selected over the remaining 11 books by the first observer was equal to the probability that Book K would be selected by the same observer in the same trial. Furthermore, the probability that book A would be selected over the remaining 11 books by the first observer was equal to the probability that Book A would be selected by the 57th observer. Thus, under the null hypothesis of random ranking, the probability of a random observer selecting one of the twelve books over the remaining 11 is equal to 1/12, or 0.083.

Customer Perception Responses

Customer Perception responses were generated from the ranking data of individual observers. However, the ranking data from all observers had to be combined for each image and each press. The most pragmatic method for combining data from many observers is to simply average the ranks for each book, recording variability statistics as well. Statistical tests for comparing population means are straight forward. These tests include *t*-tests for comparing two means and ANOVA tests for comparing multiple means. However, rank data is discrete and finite, and therefore, cannot be analyzed using methods assuming normality. Therefore, a method was required to transform the ranks from each observer to a continuous, non-finite form. Several methods are available for transforming rank data to cumulative sums that can then be transformed to standard normal scores using the inverse normal distributions. Viggiano (J. Viggiano, personal communication, March 2011) suggested using the Hazen method for transforming observer ranks to cumulative sums, expressed by Equation 15. The Hazen method was developed for use in water quality analysis applications and was proven to coincide well with parametric methods used in that field (Hunter, 2002). Thus, rank data from each observer for the Color Rendering Quality and Surface Appearance Quality experiments were transformed to Hazen cumulative probabilities using Equation 15,

$$\Phi(u) = F_{hazen}(x_i) = \frac{n + \frac{1}{2} - x_i}{n} \tag{15}$$

where n is the number of samples, 12 in this case, and x_i is the i^{th} rank. Equation 16 shows the Hazen cumulative probabilities for a set of 12 ranks.

$$\begin{bmatrix} F(12^{th}) = \frac{1}{24} \\ F(11^{th}) = \frac{3}{24} \\ F(10^{th}) = \frac{5}{24} \\ \vdots \\ F(3^{rd}) = \frac{19}{24} \\ F(2^{nd}) = \frac{21}{24} \\ F(1^{st}) = \frac{23}{24} \end{bmatrix} \tag{16}$$

Finally, the Hazen cumulative probabilities were transformed to standard normal scores—z-scores—using Equation 17.

$$u = \Phi^{-1}(F_{hazen}(x_i)) = \Phi^{-1}\left(\frac{n + \frac{1}{2} - x_i}{n}\right) \tag{17}$$

The standard normal scores were averaged for each book across observers for each image and press combination. Thus, eight sets of mean standard normal scores (four images printed on two presses) were collected for the Color Rendering Quality experiment and for the Surface Appearance Quality experiment. The data was tested for dependence on image and press before being averaged again across image and press for the final analysis.

Customer Image Quality Rating and Preference Responses

It was necessary to maintain the continuous set of data collected from each observer in the Color Rendering Quality and Surface Appearance Quality experiments because they were analyzing psychological perceptions. Customer Image Quality Rating, on the other hand, was a description of what an observer liked and did not like. Therefore, it may be said that the books an observer ranked number six or number seven were less important than the books they ranked first, second, or third. The first three rank positions provide a more relevant picture of what an observer likes, rather than what they 'kind-of' like. In a marketing application, specifically, the focus would be on creating products that an observer likes based upon studies examining those factors. In this research, the focus was on understanding what paper properties influenced an observer's decision to use a particular paper for a book of reproduced fine art. Those papers they chose not to use were of less significance. Therefore, a different approach was used to analyze the ranking data from the Image Quality experiment. Instead of transforming the rank data to a standard normal distribution, the number of times a particular book was selected in the first, second, or third ranking position was tallied. The final data set contained count data for each book, subdivided by image and press.

Count data of this nature can be described by the binomial distribution. For each observer in each trial a book could have only two outcomes: either it was selected in the top three or it was not. While administering the experiment, this author noticed observers were generally using two methods to rank the images. For the first method, observers did not rearrange the books, meaning the presentation order was still random, and simply handed the books individually to the experimenter who then recorded the rank. For the second method, observers did rearrange the books such that they were in order of their judgment. Often, observers decided upon the nine lowest-ranked books before choosing the rank of the top three.

The studentizing of binomial data requires knowledge of the probability a book is selected in the top three. This probability is expressed by Equation 18.

$$P(1^{st}\ or\ 2^{nd}\ or\ 3^{rd}) = P(1^{st}) + P(2^{nd}) + P(3^{rd}) \tag{18}$$

While the latter equation is simple, it required an understanding of the probability that a book would be selected first, second, and third. The latter probabilities were calculated for the two ranking methods described above. Figure 32 shows a Baysian tree diagram depicting the probabilities that a book is selected first, second, and third using the first ranking method. The probabilities of selecting a book first, second, third are equal, of in any position, are equal.

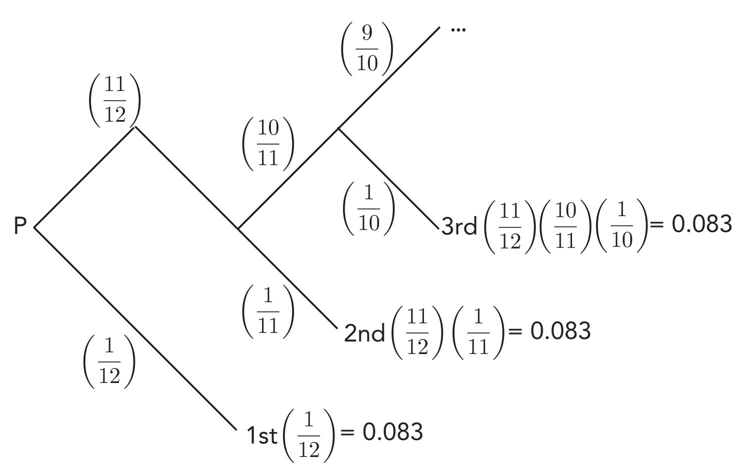


Figure 32. Probability of selection of selecting a book in any of the first three positions under the hypothesis of random selection

Figure 33 shows a Baysian tree diagram depicting the probabilities that a book is selected first, second, and third using the second ranking method. The figure shows, using the second method as well, that the probabilities of selecting a book first, second, or third are equal. While it may seem intuitive that such probabilities are equal, it was necessary to prove so using these diagrams.

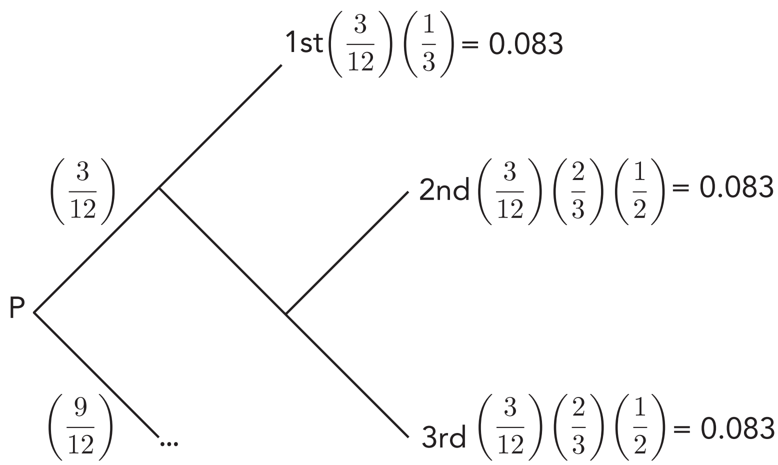


Figure 33. Probability of selecting a book in any of the first three positions after previously eliminating the bottom nine books under the hypothesis of random selection

Binomially distributed data may approximate a normal distribution should the constraints in Equation 19 be met,

$$\begin{gathered} P(X \leq x) = B(x; n, p) \approx N(np, np(1-p)) \\ with \\ np \geq 10 \ and \ n(p-1) \geq 10 \end{gathered} \quad (19)$$

where n is the number of observations, in this case 58 observers for each press, and p is the probability a book will be selected either first, second or third, in this case 0.25. The count data collected in this experiment did, in fact, satisfy the constraints for approximating a normal distribution. Therefore, Equation 20 was used to transform the Image Quality count data to standard normal scores.

$$Count_{stand} = \frac{Count - np}{\sqrt{np(1-p)}} \quad (20)$$

The same method as used to analyze the Image Quality data was used to analyze the Image Preference data. However, the Image Preference data did not contain subdivisions for the four images, but were separated based upon press.

Physical Image Parameter Selection

The PIPs for the models based upon the designed experiment were pre-defined and fixed. However, the PIPs for the expanded models were chosen to cover the gamut of paper and print quality metrics. These PIPs, other than the sample selection parameters, included caliper, mottle, dot circularity, line raggedness, print gloss, gamut volume, print contrast, and solid ink density. While it would be simple to examine a multiple regression of the complete set of 12 PIPs against the two Customer Perceptions, it was hypothesized that several of the PIPs were linearly dependent. Linear dependence among PIPs is known as multicollinearity, and can be problematic for determining least squares solutions to multiple linear regressions. Problems may include "[l]arge variances and covariances" and "[l]east squares estimates that are too large" (Montgomery, Peck, & Vining, 2006, p. 327).[18]

While it was understood that multicollinearity might have existed among the four design parameters, they were allowed to remain in the model to preserve the integrity of the design. Several methods are available for diagnosing multicollinearity among PIPs, including analysis of variance inflation factors (VIF), eigensystem analysis of **X'X** (where **X** is the regression matrix), determinant analysis of **X'X**, examination of the correlation matrix, and simple graphical examination (Montgomery et al., 2006). Paper and print quality metrics were never designed to produce linearly independent data, only to provide the industry with standardized and pragmatic quality control metric enabling the development of paper and print products. Thus, many of the tests for multicollinearity would suggest that multicollinearity was present regardless of the parameters chosen. Thus, the goal in this experiment was simply to minimize the effects of multicollinearity between PIPs under the conditions that all four sample selection parameters be included in the final set, as well as at least one PIP from the expanded set of paper and print quality metrics. Multicollinearity was analyzed using a qualitative graphical analysis and an analysis of significant correlations between PIPs. A multivariate scatter-plot between the 12 PIPs is shown in Figure 34.

Those plots circled in red were determined, both visually and by having significant correlation coefficients, to be collinear. Those PIPs with rows and columns shaded in gray were removed from the final PIP set due to the presence of significant multicollinearity. Those PIPs not shaded were included in the final expanded model. A multivariate scatter-plot of the final PIP set included in the expanded models is shown in Figure 35. Those PIPs included coolness, PST, roughness, gloss, mottle and line raggedness. While multicollinearity between the PIPs may still be present, it has been sufficiently minimized for analysis.

18 - For more information about multicollinearity, please see this reference.

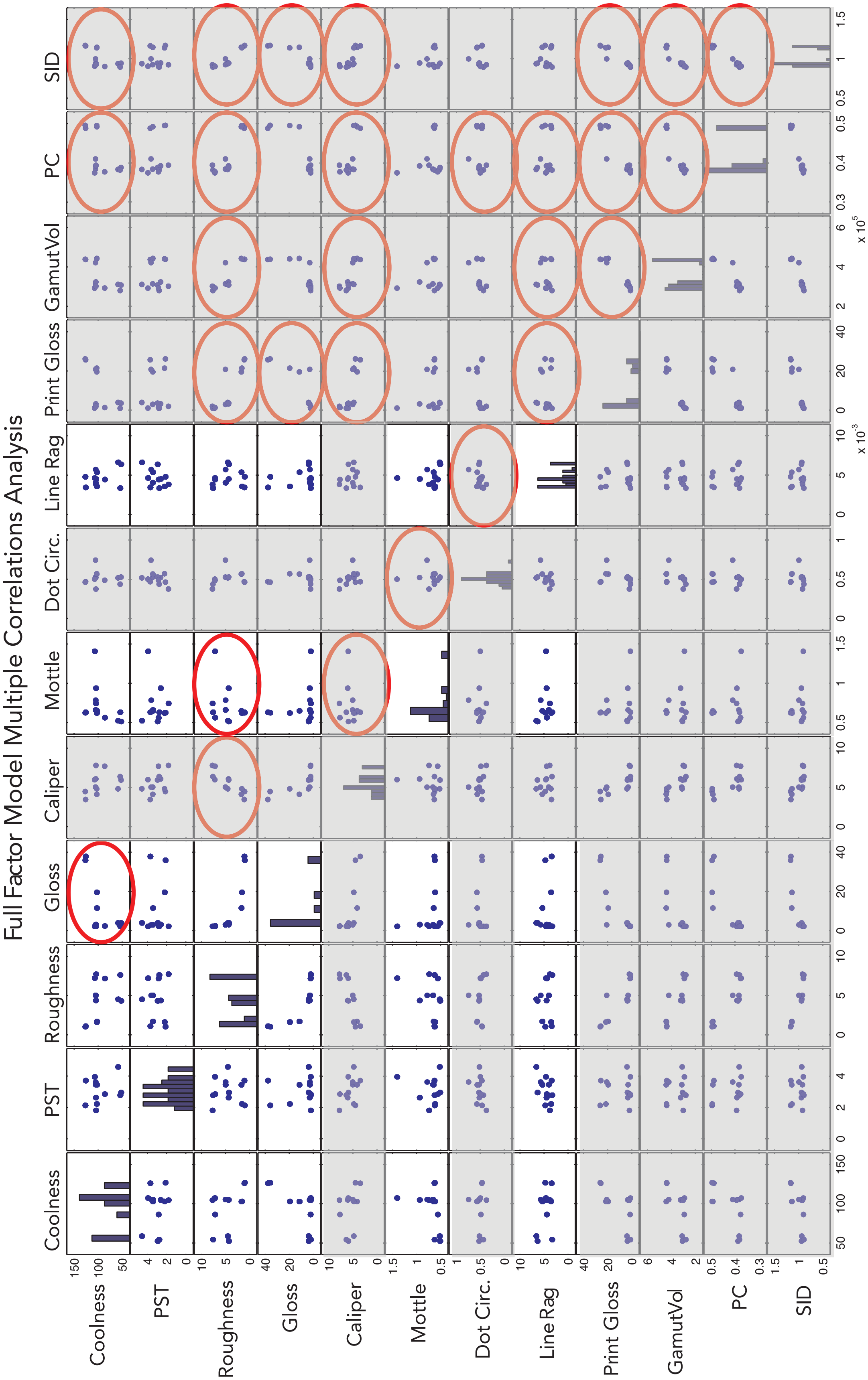


Figure 34. Matrix plot illustrating multicolinearity between measured paper and print quality metrics

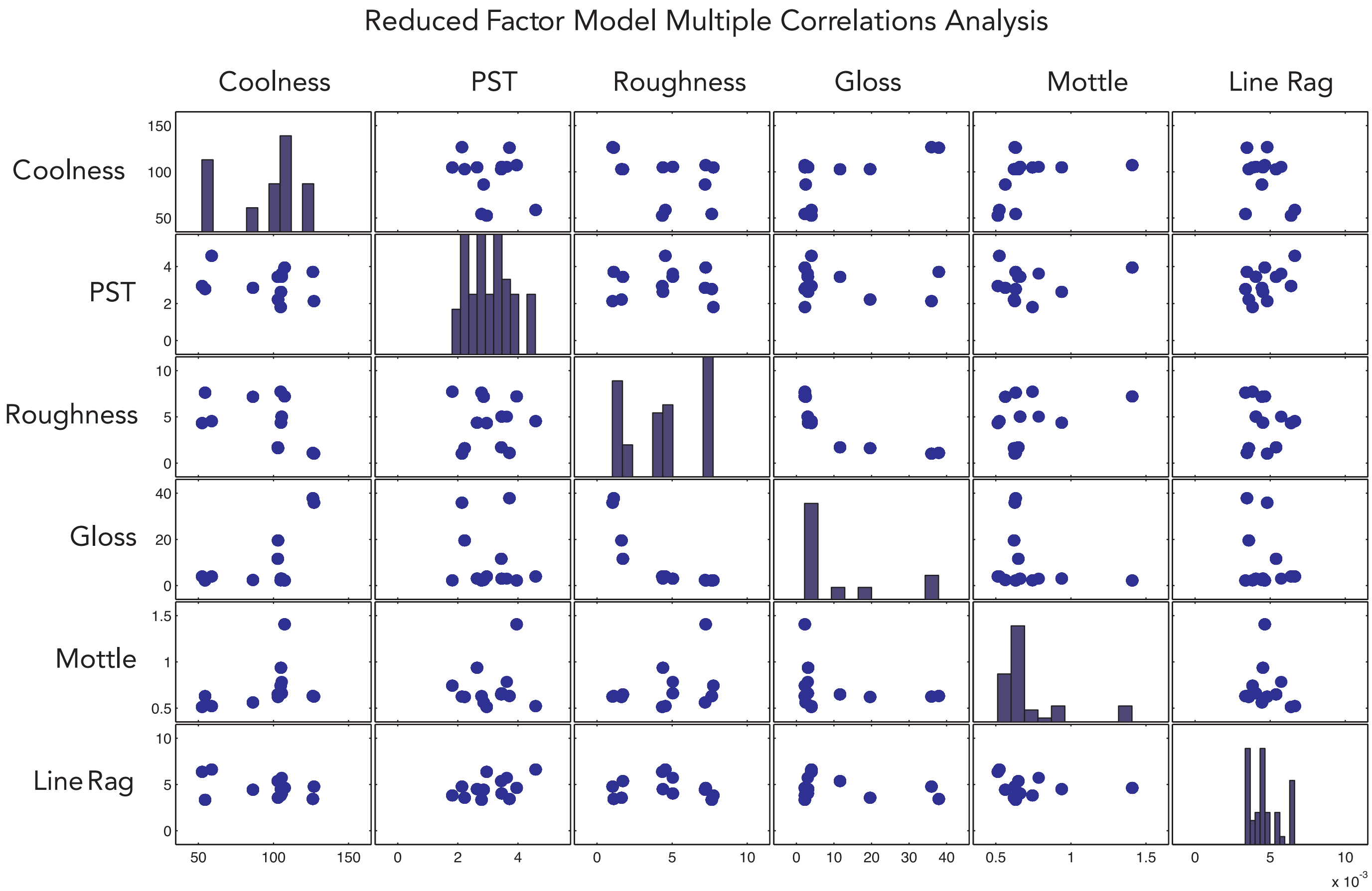


Figure 35. Matrix plot of the final reduced set of PIPs

The final set of PIPs for both the designed and expanded models have thus been determined. The following section will discuss analysis of the models considered for the Visual Algorithm component of the Image Quality Circle.

Visual Algorithms

Four Visual Algorithm models, summarized in Table 14, were examined. The psychophysical experiment was designed based upon the hypothesis that Color Rendering Quality and Surface Appearance Quality would illicit independent observer responses, and thus, could be used in a multiple regression to predict Image Quality responses. The satisfaction of this hypothesis will be discussed in later sections. In this section, multiple linear regressions for the four Visual Algorithm models—DCQ, DSQ, ECQ, and ESQ—were constructed and are analyzed.

Designed Model Analysis

The Image Quality Circle upon which the DCQ and DSQ models were based is shown in Figure 36. The Technology Variables component is shown for reference, but was not included in this experiment. The Physical Image Parameters include coolness, PST, roughness, and gloss, and the Customer Perceptions included both Color Rendering Quality and Surface Appearance Quality, although they will be analyzed separately.

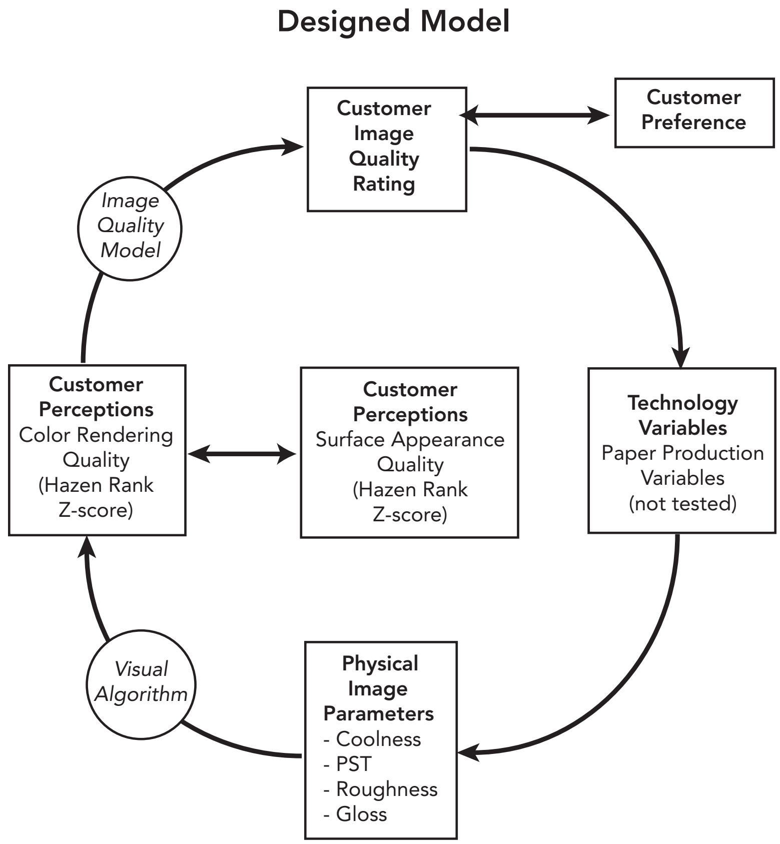


Figure 36. Image Quality Circle for the DCQ and DSQ models

Plots of the four sample selection parameters against the Color Rendering Quality Hazen z-scores are shown in Figure 37. The MATLAB™ function *stepwisefit* was used to run a stepwise linear regression for the four PIPs to determine those PIPs that, in linear combination, provided the best fit to the Color Rendering Quality data. Coolness, roughness, and gloss were determined to provide the best linear fit. The results of the linear regression are shown in Table 16. The regression coefficients, **B**, *p*-values, and whether or not the PIP was included in the model, are shown. An "In Model" parameter equal to one indicated the PIP was included in the model. Those PIPs included in the model are shown in red in Figure 37.

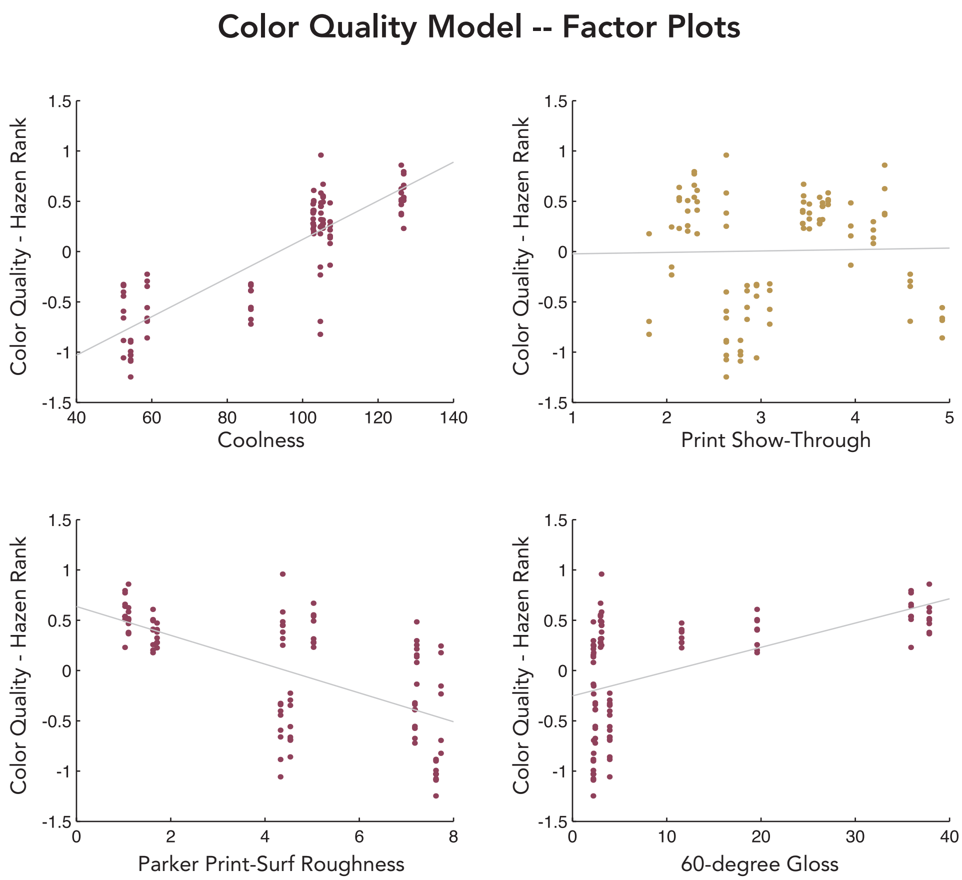


Figure 37. Factor plots for four Physical Image Parameters included in the DCQ model

Table 16. Regression statistics for the DCQ and DSQ models[19]

DCQ					
	Intercept	Coolness	PST	Roughness	Gloss
B	-0.979	0.02	0.036	-0.141	-0.022
p-val	--	< 0.001	0.201	< 0.001	< 0.001
In Model	1	1	0	1	1
DSQ					
	Intercept	Coolness	PST	Roughness	Gloss
B	0.09	0.004	-0.047	-0.081	-0.009
p-val	--	0.003	0.159	< 0.001	0.027
In Model	1	1	0	1	1

The standardized residuals were analyzed to determine the success of the DCQ linear regression. The residual analysis plots are shown in Figure 38. The normal probability plot of standardized residuals, shown in Figure 38a, suggests that the residuals are normally distributed, with the exception of two possible outliers. Figure 38b shows the

19 - 'B' defines the regression coefficients, 'p-val' defines the significance of the model parameter, and 'In Model' defines whether or not the parameter was included in the model based on its statistical significance.

standardized residuals plotted against fitted values, and suggests there is continuity of variance among the residuals. Figure 38c shows the standardized residuals plotted against order number. Most of the variation appears random, with the exception of the potential outliers with standardized residual values less than -3. However, it is necessary to note that two outliers were previously removed from the data set. The images associated with those outliers, Paper D with the Oil Painting from both presses, accumulated significant damage due to abrasion throughout the course of the study. This negatively impacted image quality. In addition, the effects of abrasion were noted by several observers. Some understood the damage to be caused by abrasion and others judged the image poorly without realizing the problem was due to damage, as opposed to poor printing quality. Therefore, while two more outliers became present after the removal of the latter two, this author felt that any further removal of data would distort the results, especially since so few papers and images were included from the beginning.

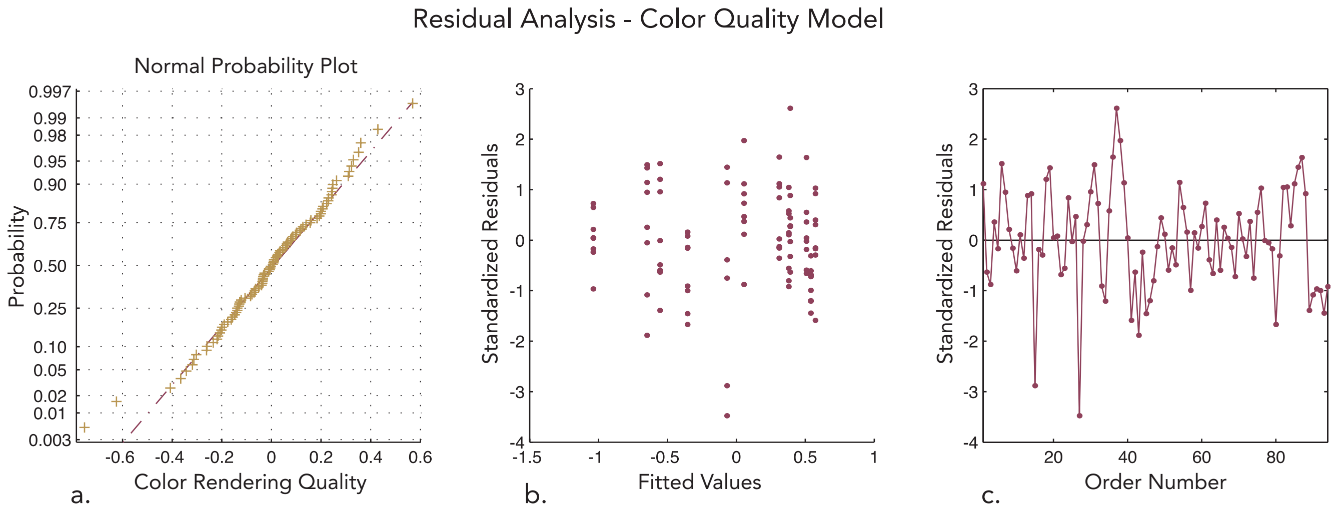


Figure 38. Residual plots for the DCQ model

Plots of the four sample selection parameters against the Surface Appearance Quality Hazen z-scores are shown in Figure 39. As for the Color Rendering Quality analysis, coolness, roughness, and gloss were determined to be those PIPs that result in the most accurate fit to the Surface Appearance Quality data. The regression statistics for this fit are shown in Table 16. Those factors included in the model are shown in red in Figure 39.

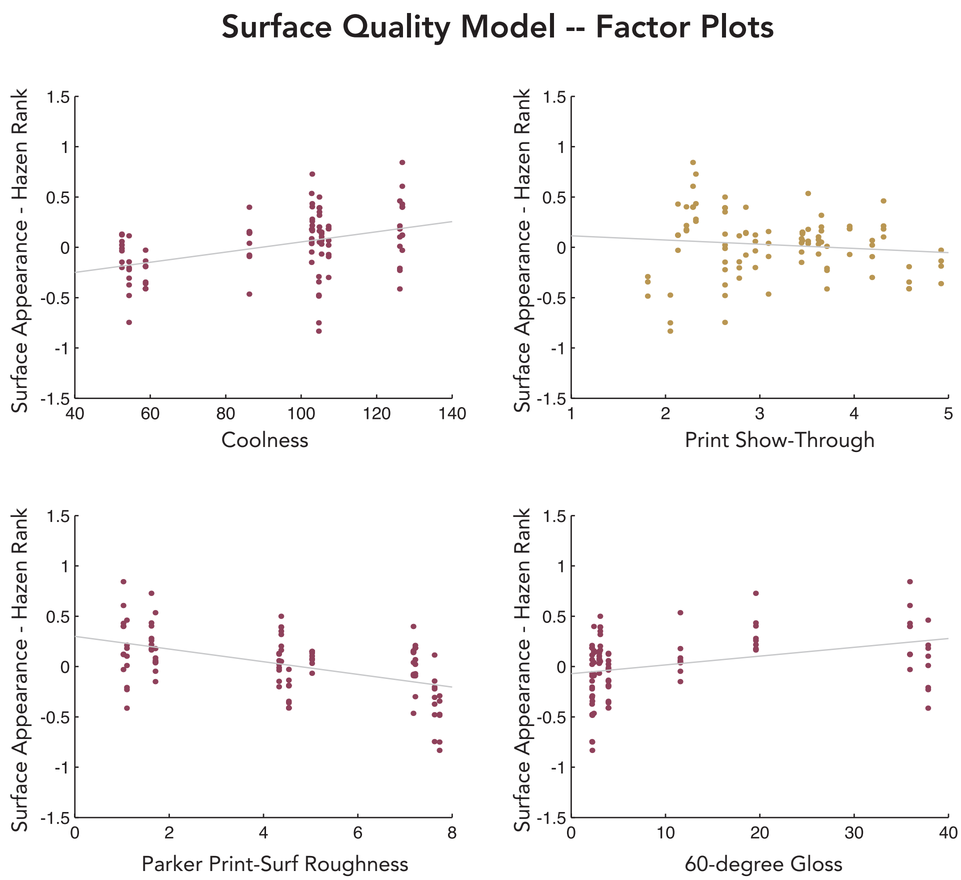


Figure 39. Factor plots for four Physical Image Parameters included in the DSQ model

Standardized residual plots for the DSQ model are shown in Figure 40. The DSQ residual plots suggest that the data satisfy the requirements for analysis using the normal distribution, that the residuals are normally distributed (Figure 40a), of equal variance, (Figure 40b), and random with respect to order, without any definite outliers (Figure 40c).

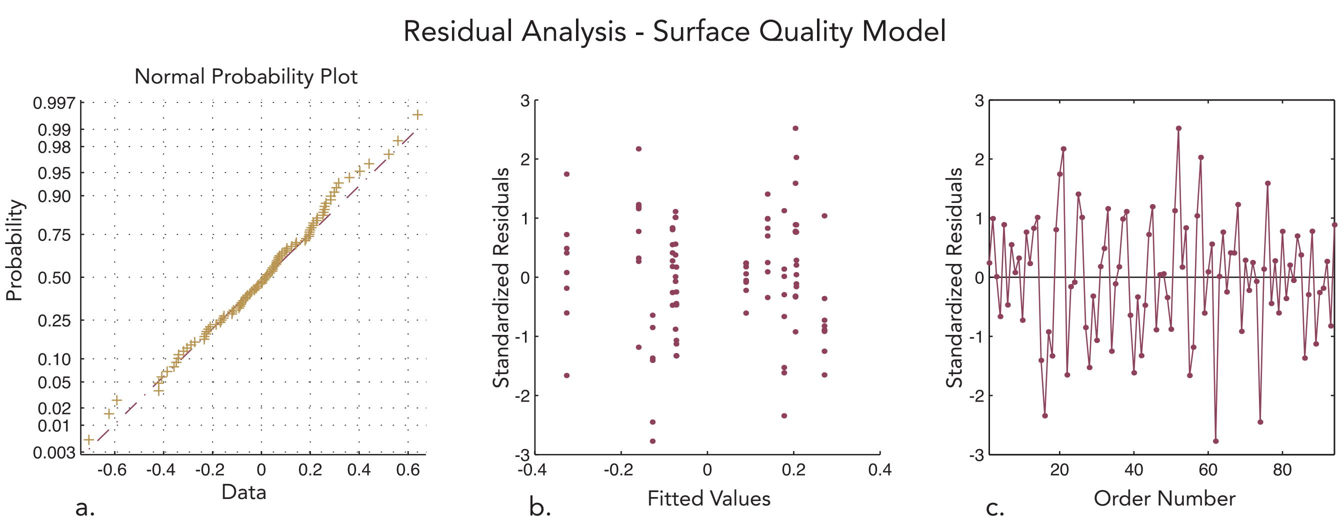


Figure 40. Residual plots for the DSQ model

While linear regression models satisfying the requirements for analysis using the normal distribution were developed for both the DCQ and DSQ data, there is no guarantee that both models will be used to develop the final Visual Algorithm. Such an analysis will be discussed in later sections. The following sections will discuss the analysis of linear regression models for the ECQ and ESQ models.

Expanded Model Analysis

The Image Quality Circle upon which the ECQ and ESQ models were based is shown in Figure 41. As in Figure 36, the Technology Variables component is shown only for reference. The PIPs were expanded to include 40% Mottle and Line Raggedness.

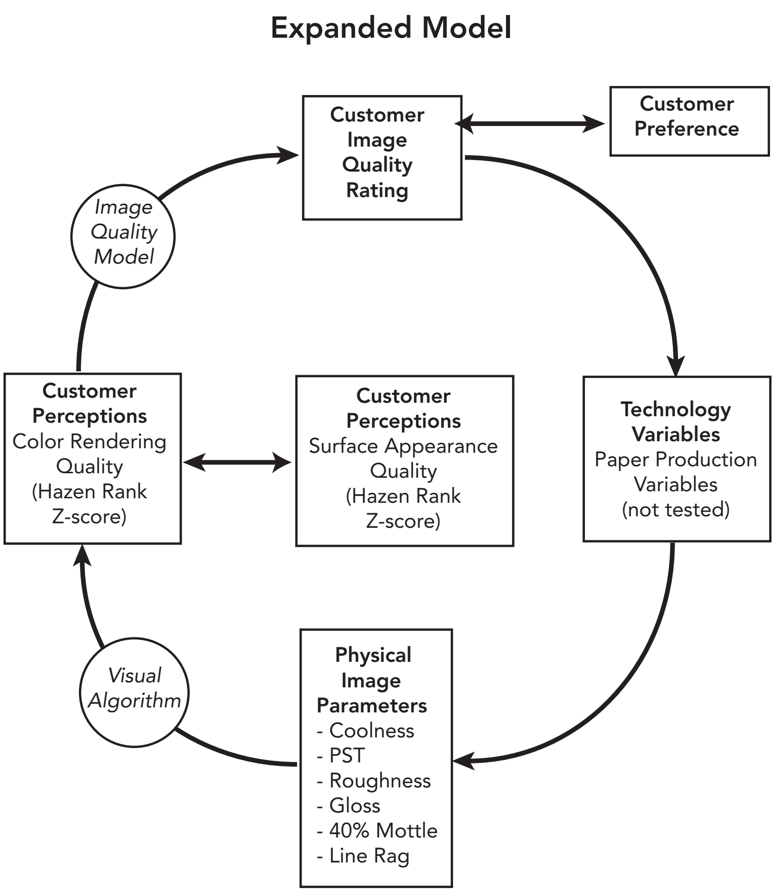


Figure 41. Image Quality Circle for the ECQ and ESQ models

Analysis of the expanded models was conducted in the same manner as for the designed models. Color Rendering Quality and Surface Appearance Quality were the responses in two separate regressions. The PIPs included the four sample selection parameters with the addition of 40% Mottle and Line Raggedness (Line Rag). The MATLAB® function *stepwisefit* was used to run a stepwise regression for determining which of the six PIPs provided the best linear fit to the responses. The regression statistics are shown in Table 17. While the expanded models did include two additional parameters compared to the designed models, the predictors resulting in the smallest *adjusted* R^2 were once again coolness, roughness, and gloss.

Table 17. Regression statistics for the ECQ and ESQ models

ECQ							
	Intercept	Coolness	PST	Roughness	Gloss	40% Mottle	Line Rag.
B	-0.979	0.02	0.036	-0.141	-0.022	0.140	-17.231
p-val	--	< 0.001	0.201	< 0.001	< 0.001	0.223	0.274
In Model	1	1	0	1	1	0	0
ESQ							
	Intercept	Coolness	PST	Roughness	Gloss	40% Mottle	Line Rag.
B	0.09	0.004	-0.047	-0.081	-0.009	-0.114	-18.783
p-val	--	0.003	0.159	< 0.001	0.027	0.401	0.310
In Model	1	1	0	1	1	0	0

Figure 42 shows plots of the six PIPs against Color Rendering Quality. The three PIPs included in the regression are shown in red.

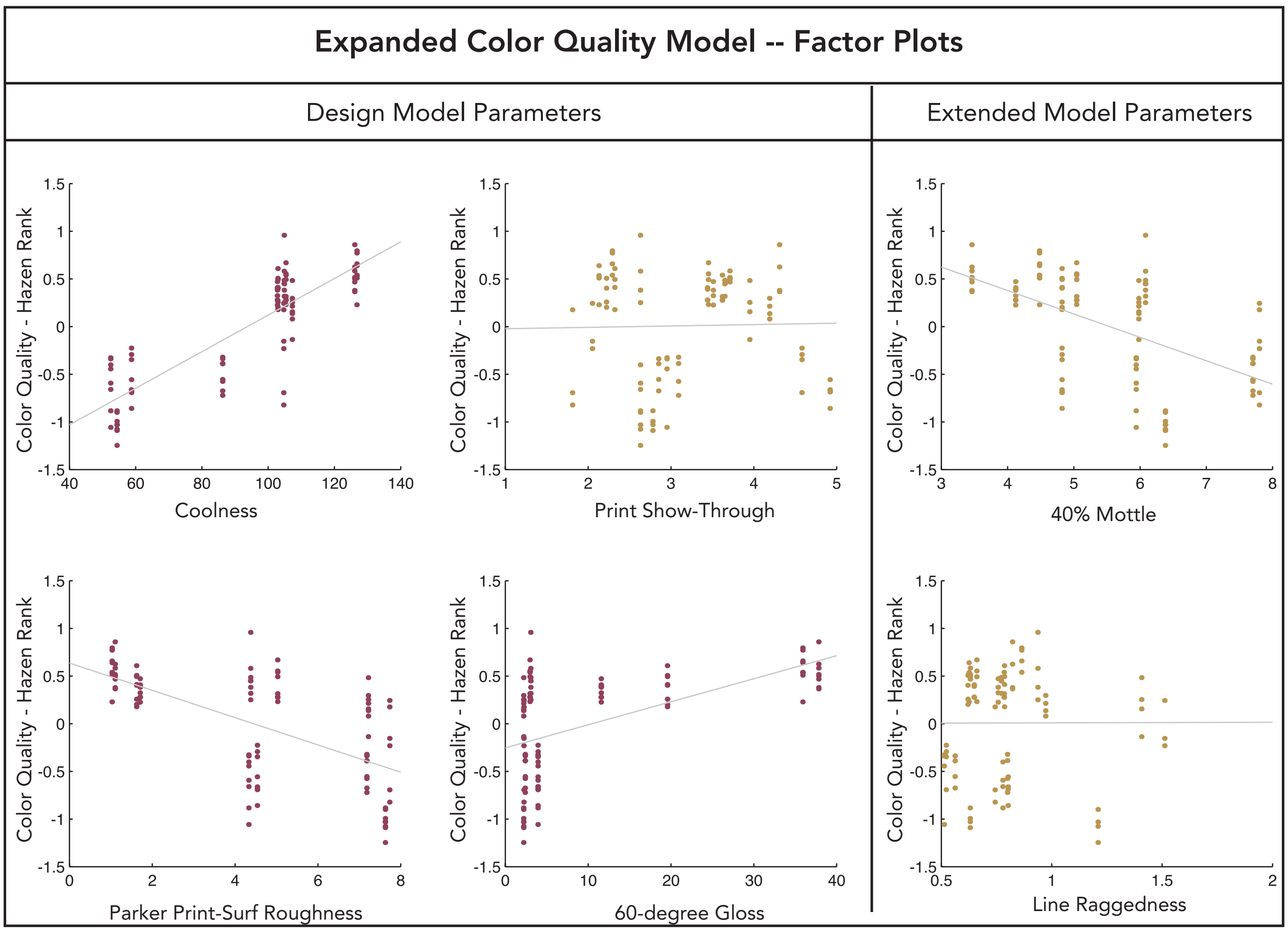


Figure 42. Factor plots for four Physical Image Parameters included in the ECQ model

Similarly, Figure 43 shows plots of the six PIPs against Surface Appearance Quality. With exception of the additional two parameters, the plots in Figures 42 and 43 are the same as those in Figures 37 and 39.

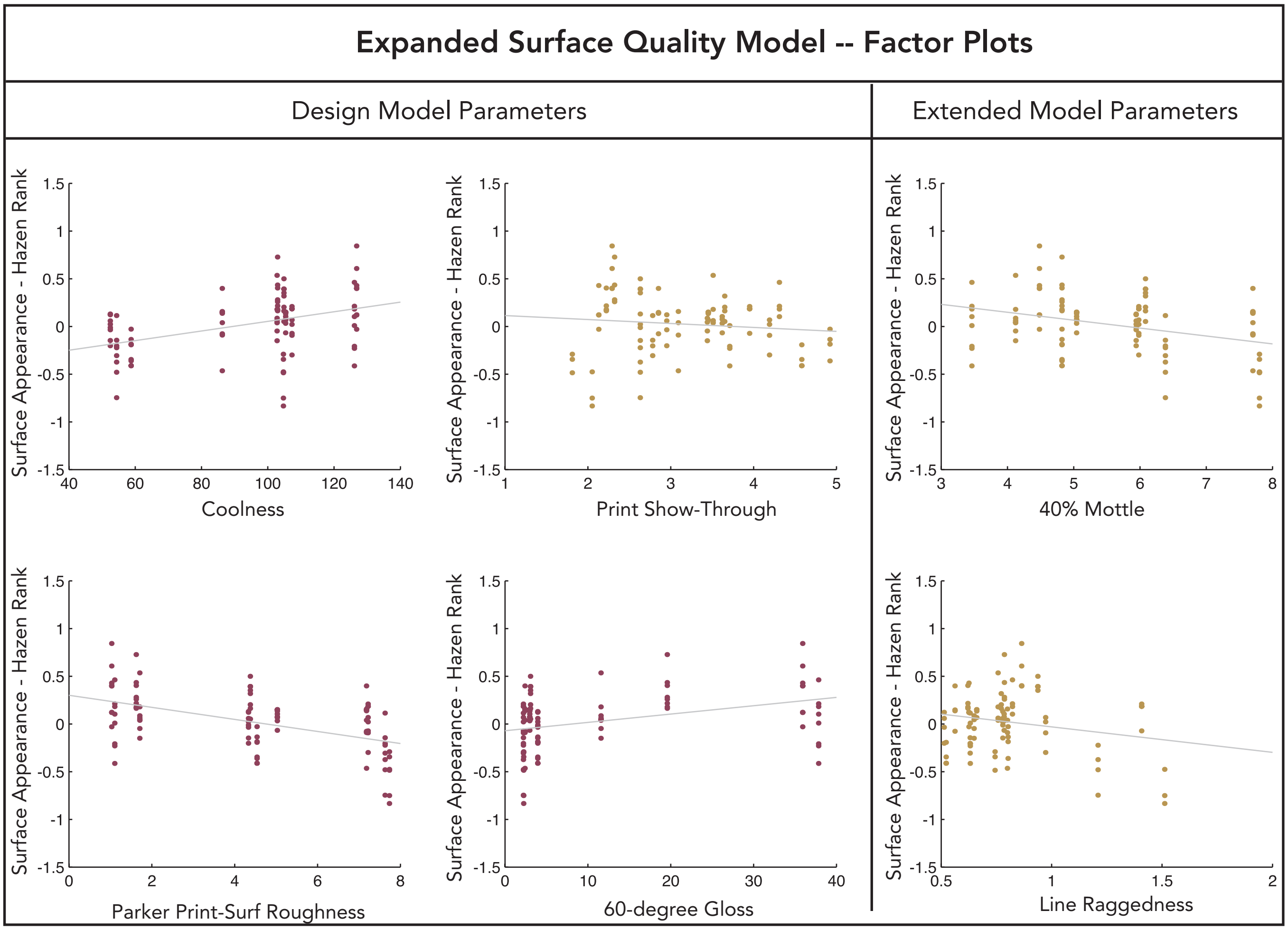


Figure 43. Factor plots for four Physical Image Parameters included in the ESQ model

Image Quality Model

According the Image Quality Circle, Customer Image Quality Rating is predicted directly from Customer Perceptions. The model used to make that prediction is called an Image Quality Model. Customer Image Quality Rating was measured by the Image Quality stage of the psychophysical experiment. Customer Image Quality Rating may be interpreted as a measure of behavior, describing choices people make based upon likes and dislikes, while Customer Perceptions are measurements of more subconscious elements of decision-making that force observers to focus on elements of decision-making of which they would not normally be attentive. However, the statistical construction of the Image Quality Models had the same problems of multicollinearity as the Visual Algorithms. The development of the DIQ and EIQ models are described below. Each began with an analysis of multicollinearity followed by the fitting of a linear regression model.

Designed Model

Multicollinearity was analyzed by plotting the fitted values of Color Rendering Quality against the fitted values of Surface Appearance Quality modelled by the Visual Algorithms, shown in Figure 44a. The plot suggests a strong correlation between Color Rendering Quality and Surface Appearance Quality, which indeed was the case ($r = 0.92$, $p <= 0.001$). Therefore, it was necessary to choose either Color Rendering Quality or Surface Appearance Quality for the final DIQ model. Plots of Standardized Image Quality Counts against Color Rendering Quality and Surface Appearance Quality are shown in Figures 44b and 44c, respectively.

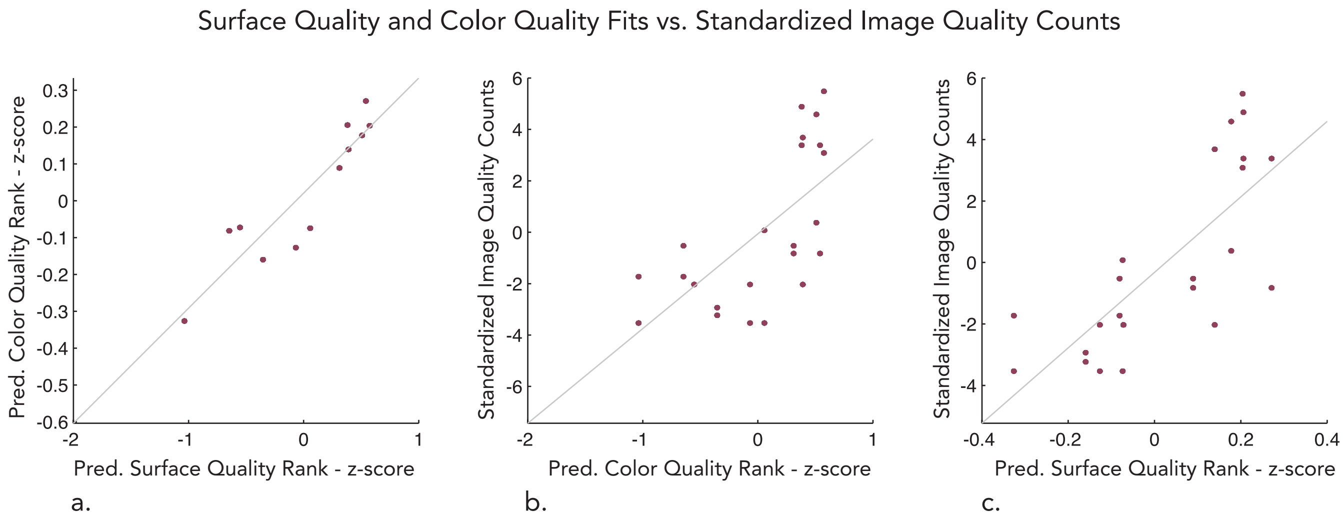


Figure 44. Plots of (a) predicted Color Quality Rank versus predicted Surface Quality Rank, (b) standardized Image Quality Counts versus predicted Color Quality Rank, and (c) standardized Image Quality Counts versus predicted Surface Quality Rank

The decision was made to use Color Rendering Quality as the predictor for Image Quality because the Color Rendering Quality was tested before Surface Appearance Quality during the psychophysical experiment. Thus, it is highly likely that observers were influenced by having previously analyzed color quality when judging surface appearance quality.

Standardized residual plots for the linear fit between Color Rendering Quality fits and Standardized Image Quality Counts are shown in Figure 45. The normal probability plot in Figure 45a suggests the residuals are not normally distributed. This may be attributed to the possibility that a linear fit was not ideal or unattributed factors were at work. However, further research must be conducted to determine whether another model would be more suitable and whether there is a psychological basis for a nonlinear model. The remaining residual plots, in Figures 45b and 45c, do not suggest a lack of equal variance or present outliers. The cyclical nature of 45c is due to the repetition of rank data across image and press.

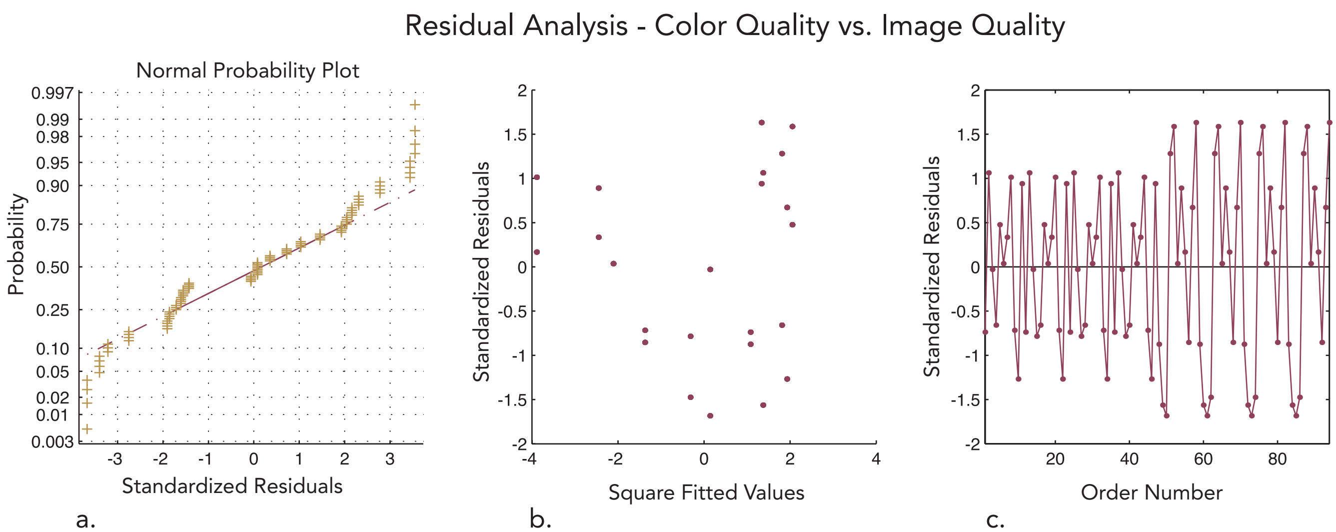


Figure 45. Residual plot of the regression between standardized Image Quality Counts and predicted Color Quality Rank

Expanded Model

The ECQ model contained the same PIPs as the DCQ model, and thus, the Color Rendering Quality fits for the designed model were identical to the expanded model. Therefore, the EIQ model is identical to the DIQ model and requires no further discussion.

Bypass Customer Perceptions

Up to this point, models have been discussed based upon a strict interpretation of the Image Quality Circle. However, there are times when short cuts must be taken. Examples of short cuts taken within the Image Quality Circle are illustrated in Farnand (2008, 2009). In fact, Engeldrum himself (1999) suggested the possibility of such short cuts: an example would be creating Image Quality Models from Physical Image Parameters if Customer Perceptions are not known or cannot be interpreted. In this experiment, models will be tested in which Customer Image Quality Rating is predicted directly from the Physical Image Parameters. Although generating these models is not the main goal of this thesis, it is important to acknowledge the possibility that they can be used. This would have especially been important if the Color Rendering Quality and Surface Appearance Quality experiments had yielded inconclusive results.

Designed Model - Bypass Customer Perceptions

The model describing the connection between Physical Image Parameters and Customer Image Quality Rating is here called the Direct Image Quality Model, although Engeldrum (1999) referred to them as the "stimulus" form of Image Quality Models. Figure 46 shows the abbreviated Image Quality Circle incorporating the Direct Image Quality Model for the designed experiment, the DD2IQ model.

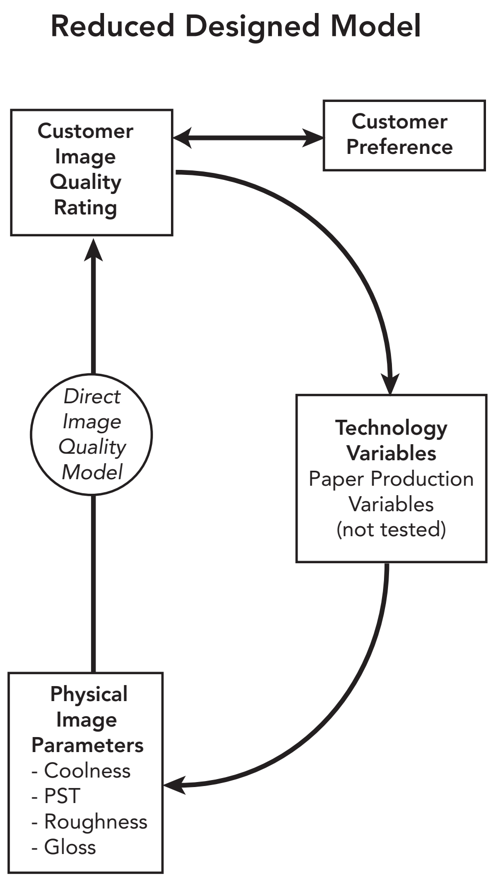


Figure 46. Image Quality Circle for DD2IQ model

The linear regression was computed using the same stepwise method used to analyze the Visual Algorithms and Image Quality Models. The regression statistics for the DD2IQ model are shown in Table 18. Coolness, PST and roughness resulted in the best linear fit to the Standardized Image Quality Count data.

Table 18. Regression statistics for the DD2IQ model

	Intercept	Coolness	PST	Roughness	Gloss
B	1.043	0.023	-0.744	-0.821	0.002
p-val	--	0.003	<0.001	< 0.001	0.932
In Model	Yes	Yes	Yes	Yes	No

Figure 47 shows plots of the four sample selection parameters against Standardized Image Quality Counts. While coolness, PST, and roughness were included in the model based on statistical significance, PST does not appear to show a significant visual trend. Further research is needed to determine whether or not PST should be included in the final DD2IQ model. For that reason, only coolness and roughness are shown in red in Figure 47.

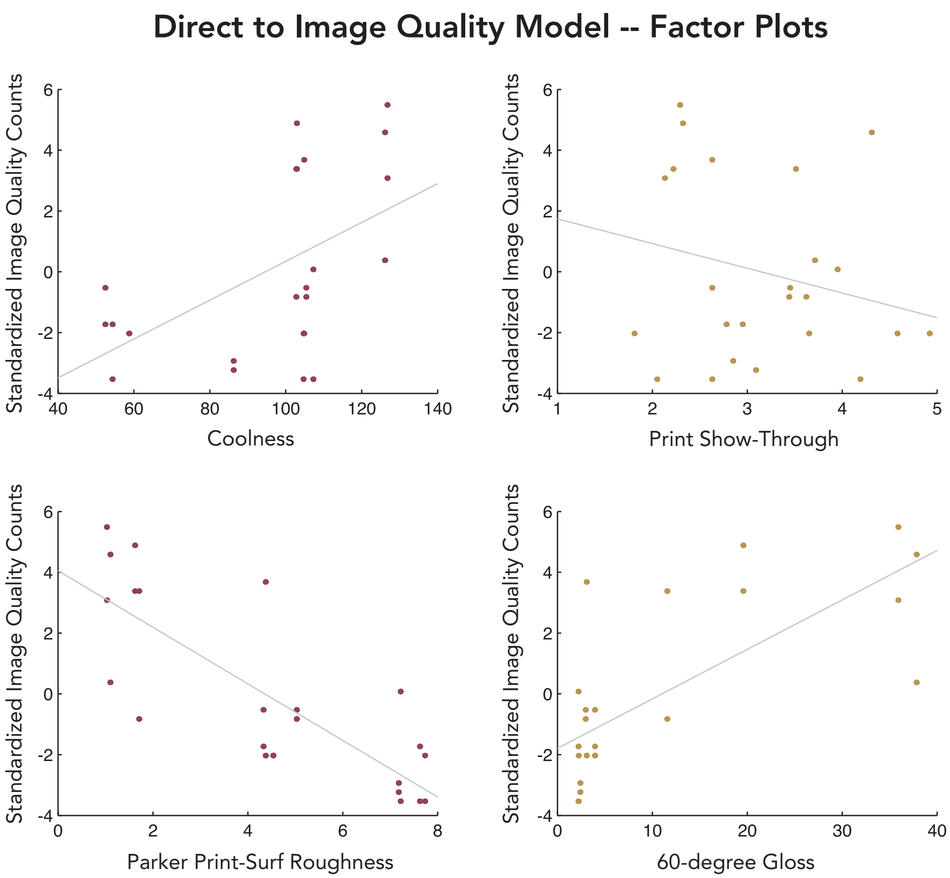


Figure 47. Factor plots for the four Physical Image Parameters in the DD2IQ model

Standardized residual plots for the DD2IQ model, including PST, are shown in Figure 48. With the exception of a few points near the ends of the normal distribution plot (Figure 48a), the standardized residuals appear to be normally distributed, to have equal variance (Figure 48b), and to be random with respect to order number (Figure 48c), with exception given to the previously explained cyclical nature of this plot.

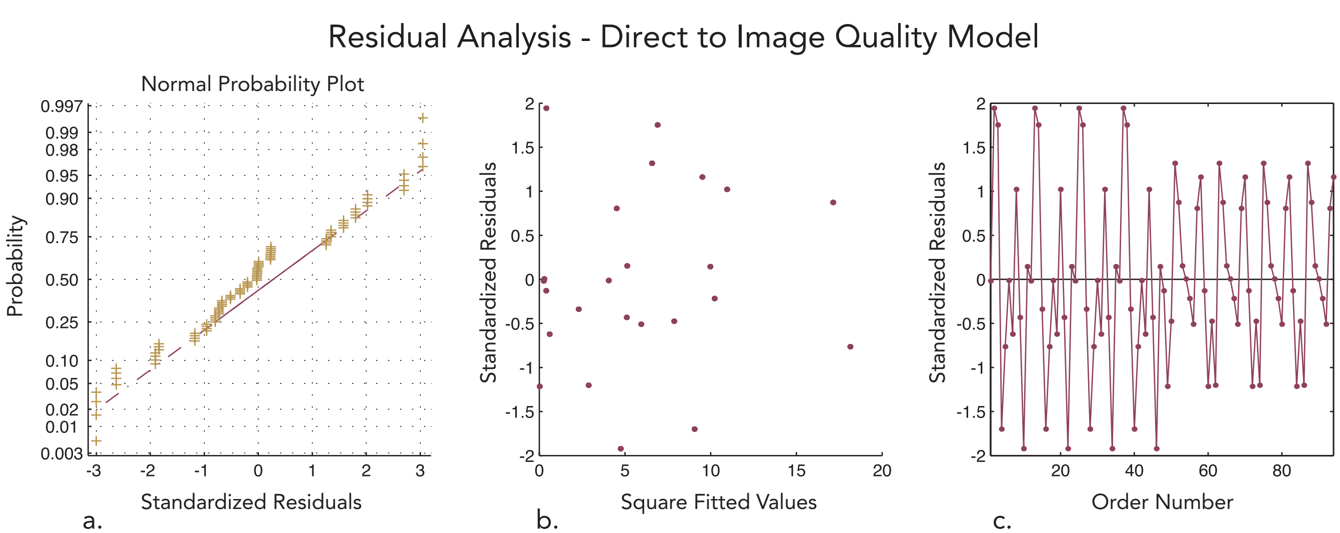


Figure 48. Residual plots for the DD2IQ model

Expanded Model - Bypass Customer Perceptions

Figure 49 shows the abbreviated Image Quality Circle incorporating the Expanded Direct Image Quality Model for the designed experiment, the ED2IQ model.

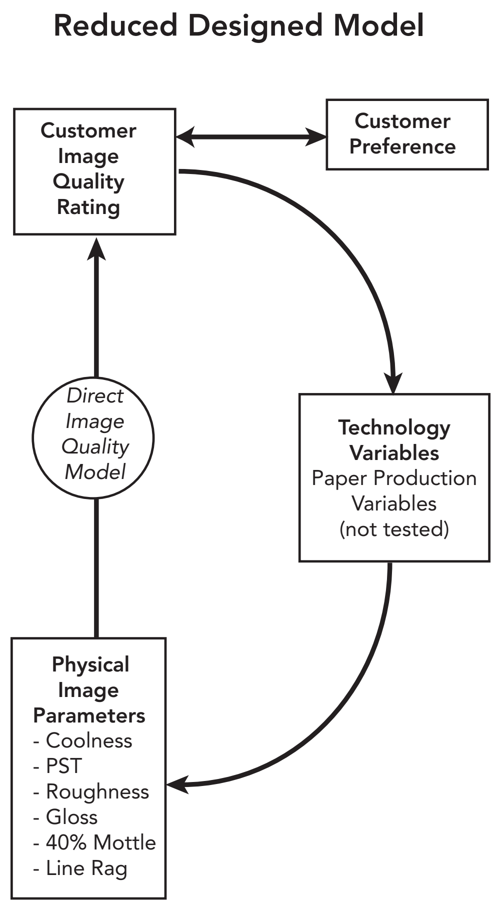


Figure 49. Image Quality Circle for the ED2IQ model

Unlike the case encountered when analyzing the Visual Algorithms, the ED2IQ model was fit using different parameters than the DD2IQ model. In this case, PST, roughness, and 40% mottle were determined to be the best model PIPs (see Table 19).

Table 19. Regression statistics for the ED2IQ model

	Intercept	Coolness	PST	Roughness	Gloss	40% Mottle	Line Rag.
B	4.717	0.014	-0.772	-1.046	0.014	2.871	-69.762
p-val	--	0.08	< 0.001	< 0.001	0.515	<0.001	0.506
In Model	Yes	No	Yes	Yes	No	Yes	No

However, upon visual analysis of the plots of Standardized Image Quality Counts against the six PIPs as shown in Figure 50, it appears that coolness also has a strong linear relationship with image quality, possibly even more so than PST. This is another finding that will require further research. The three PIPs included in the ED2IQ model, as well as coolness, are shown in red.

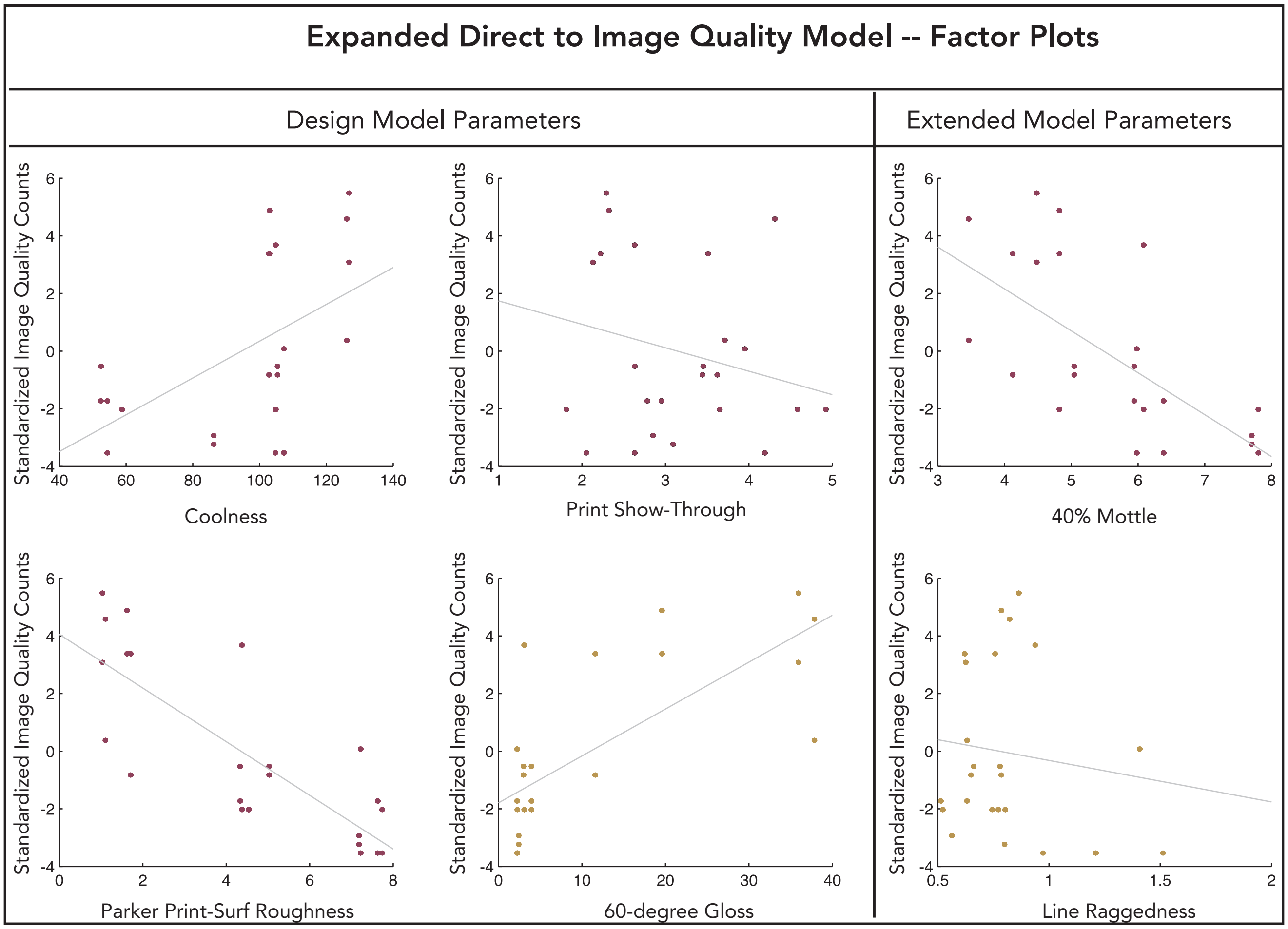


Figure 50. Factor plots for the six Physical Image Parameters in the ED2IQ model

Standardized residual plots for the ED2IQ model without coolness are shown in Figure 51. With the exception of a few points near the ends of the normal distribution plot (Figure 51a), the standardized residuals appear to be normally distributed, to have equal variance (Figure 51b), and to be random with respect to order number (Figure 51c).

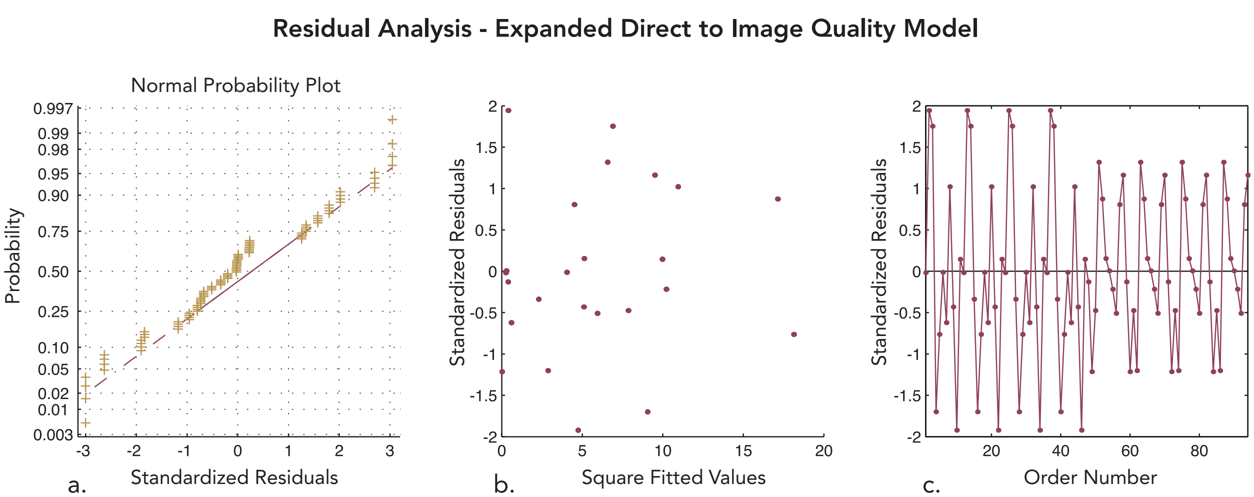


Figure 51. Residual plots for the ED2IQ model

Expanded Model - Substitutions

The linear dependence among PIPs included in the final model was previously discussed, and the author felt it important to outline one possible component of that linear dependence. In Figure 52, roughness is plotted against a linear combination of gloss and 40% mottle. Roughness is highly correlated with the linear fit of gloss and 40% mottle ($r = 0.95$, $p <= 0.001$). Thus, it is possible that, for models including roughness, roughness may be replaced by gloss and 40% mottle without significant loss in accuracy, although with an increase in model complexity. Likewise, models including both gloss and 40% mottle may be reduced to include only roughness.

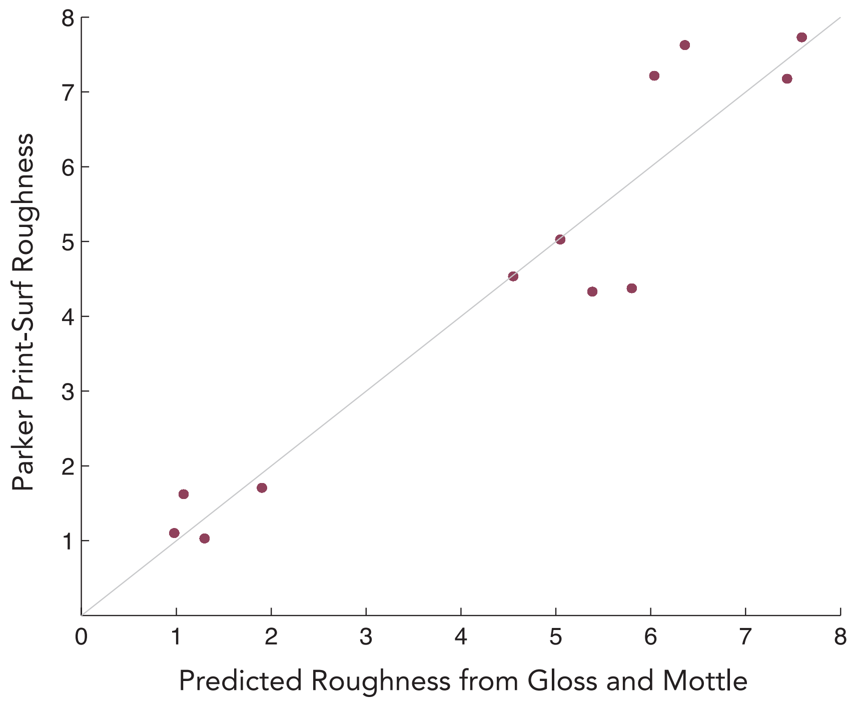


Figure 52. Regression of gloss and mottle against roughness

Complete System

The complete set of models for the Image Quality Circle was determined based upon the previously discussed analysis. The final linear models are shown in Table 20. The Visual Algorithm satisfies both the designed and expanded models because both analyses found the same PIPs significant. By the same reasoning, the Image Quality model satisfies both the designed and expanded models. However, the Direct Image Quality Model, shown in Table 20, satisfies only the designed model. The designed model was chosen here for the purposes of continuity to the psychophysical experiment, but the expanded model was more precise as illustrated by the higher *adjusted* R^2 value.

Table 20. Three final models for the complete Image Quality Circle

Model Component	Model
Visual Algorithm	CQ = −0.9791 + 0.0196*C − 0.1411*R − 0.0219*G
Image Quality Model	CIQR = −0.0611 + 3.6880*CQ
Direct Image Quality Model	CIQR = 1.043 + 0.023*C − 0.744*P −0.821*R

CQ = Color Rendering Quality, C = Coolness, R = Roughness

G = Gloss, CIQR = Customer Image Quality Rating

Contour plots are shown in Figure 53, illustrating the change in Color Rendering Quality as a function of coolness, roughness, and gloss. The response is shown as a *z*-score, and changes positively as a function of coolness, negatively as a function of roughness, and negatively as a function of gloss. Thus, a paper expected to have a high Color Rendering Quality will have high coolness, low roughness, and low gloss. By extension, Surface Appearance Quality and Customer Image Quality Rating are also optimized for papers with high coolness, low roughness, and low gloss.

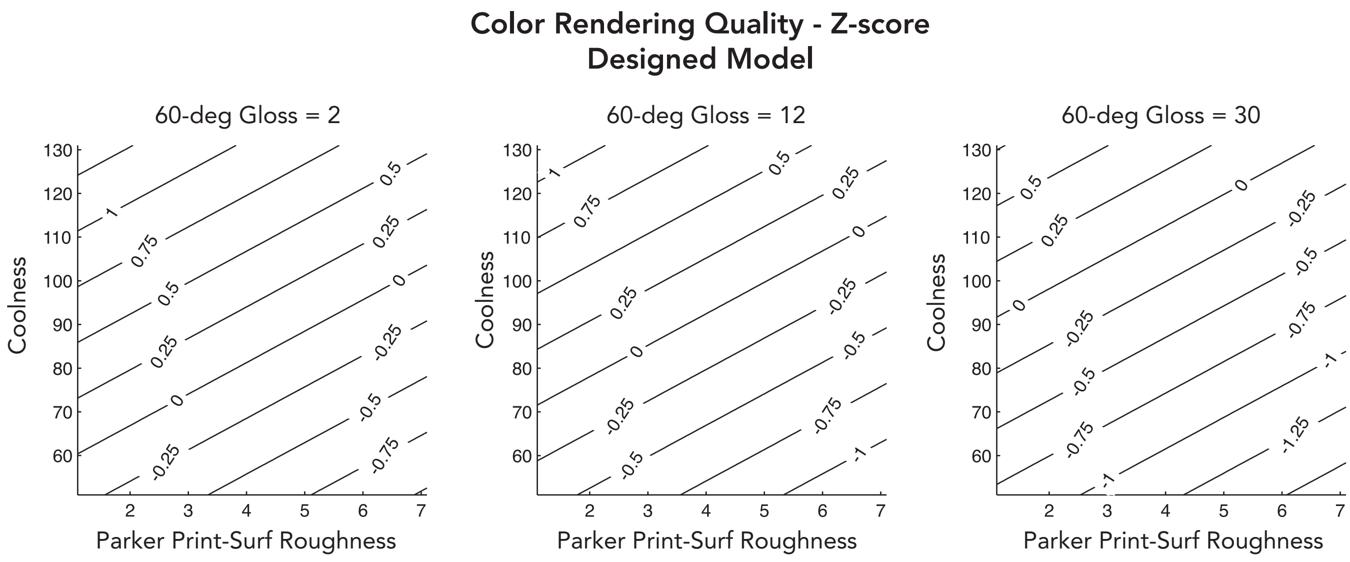


Figure 53. Contour plots showing the relationship between mottle, coolness and roughness relative to DCQ model predictions

Contour plots for the Direct Image Quality Model are shown in Figure 54. Gloss did not prove to be a significant factor in this model; however, Customer Image Quality Rating was still optimized with high coolness, low roughness, and, in this case, low PST. The low PST may also be extended to high Basis Weight because the two factors were confounded.

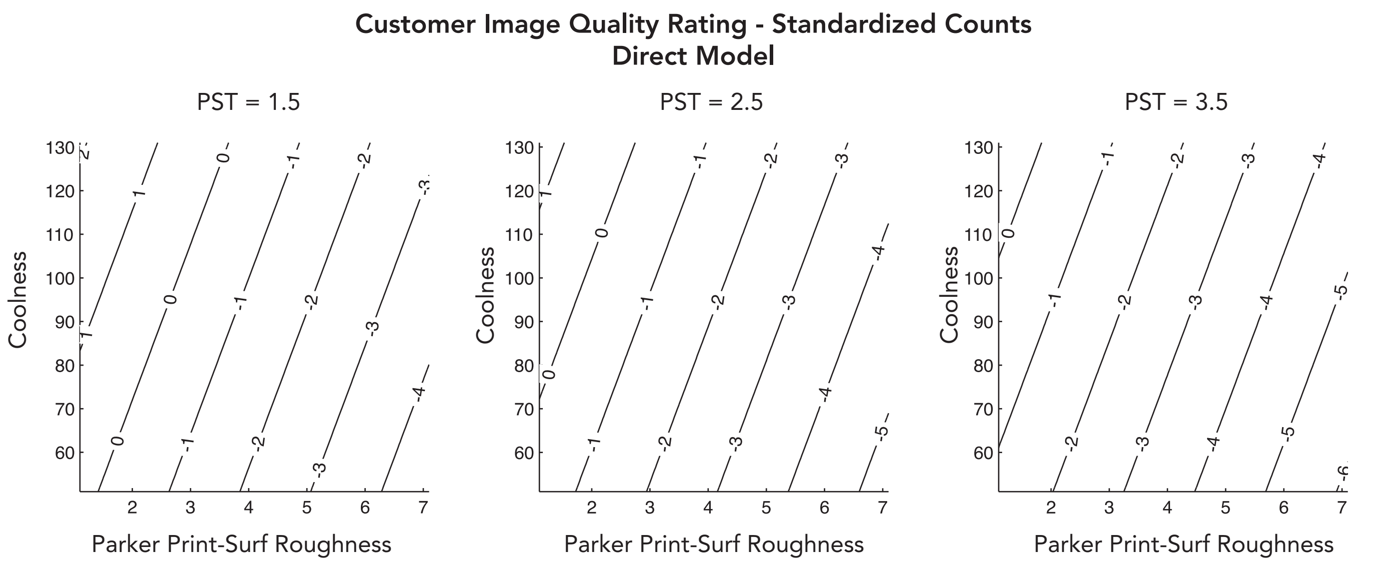


Figure 54. Contour plots showing the relationship between PST, coolness, and roughness relative to DD2IQ Model Predictions

Preference Analysis

While the Image Quality Circle did not explicitly model preference, the relationship between Customer Image Quality Rating and Customer Preference was analyzed to determine if Customer Preference could be predicted by Customer Image Quality Rating. Figure 55 depicts this relationship. There is a significant correlation between Customer Image Quality Rating and Customer Preference ($r = 0.73$, and $p <= 0.001$), suggesting a strong linear relationship.

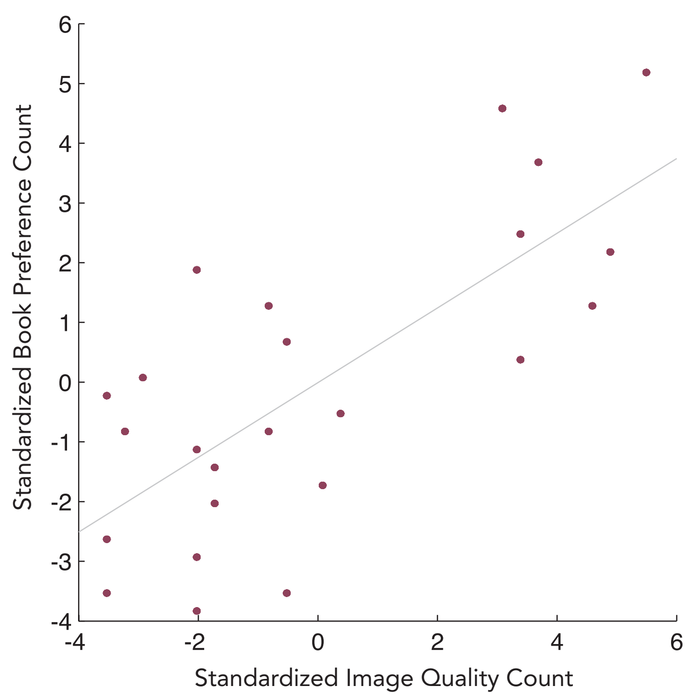


Figure 55. Scatter plot of standardized Image Quality Counts versus standardized Book Preference Counts

The standard residual plots for the regression of standardized Book Preference Count versus standardized Image Quality Count are shown in Figure 56. The plots suggest that the standardized residuals are normally distributed (Figure 56a), equivarient (Figure 56b), and random with respect to order number (Figure 56c).

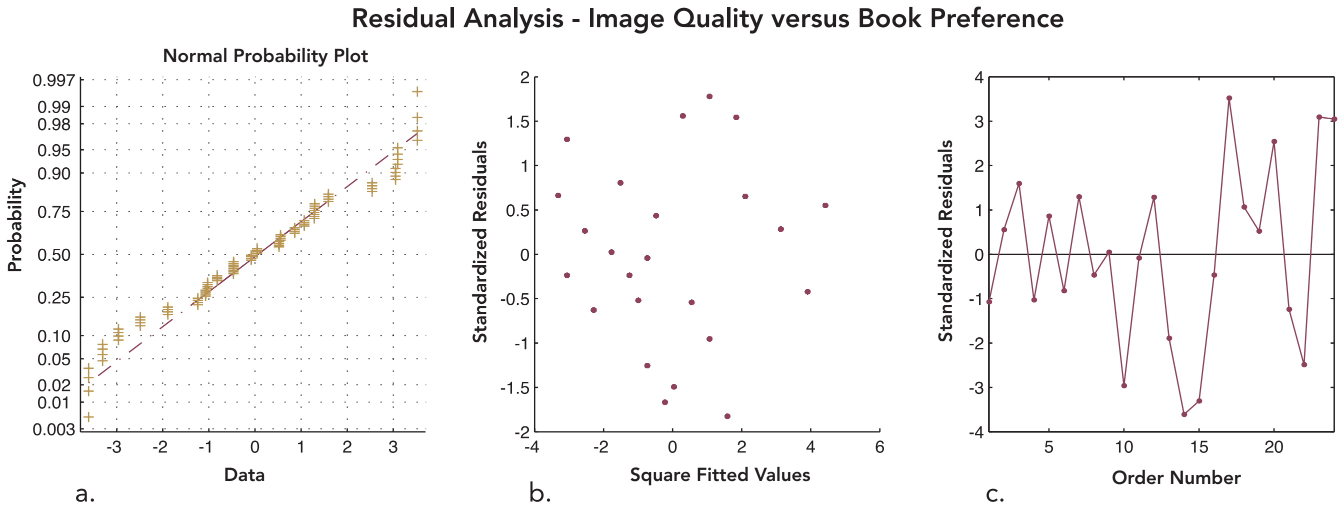


Figure 56. Residual plots for the plot of standardized Image Quality Counts versus standardized Book Preference Counts

Thus, a model of Customer Preference from Customer Image Quality Rating is the final element for the complete Image Quality Circle analyzed in this thesis. The following section will analyze responses provided by observers following each of the four experiments and the relationship between these responses and the psychophysical experiment results.

Lexical Analysis

Each observer was asked to describe the factors that influenced their judgements for each of the four psychophysical experiments (image quality, color rendering quality, surface appearance quality, and preference). The responses were collected from each observer and compiled respective to the four experiments. The observer responses were collected to support the psychophysical statistical analysis and to provide insight into observer reasoning. There are many methods for analyzing this type of natural language data. In this experiment, the frequency of which specific words or ideas were used by individuals was analyzed. This was the most efficient method for accomplishing the goals of the lexical analysis. The semantics in which the words were used varied for each participant and also varied depending on the field in which the participant worked or studied. Photographers tended to reason differently from scientists and engineers, who also reasoned differently from designers. Figure 57 illustrates the variety of words used by participants to describe their decision-making reasoning following the Color Rendering Quality Experiment. The collective words are grouped into eight general categories: words related to color, saturation, image detail, tone, memory, gloss, paper, and general preference. Those words highlighted in yellow were the ten most frequently used words for all participants. Some words were generalized to reduce redundancy. For example, 'cool' and 'coolness' were counted as the same word, and 'too light' and 'lighter' were generalized to 'lightness.'

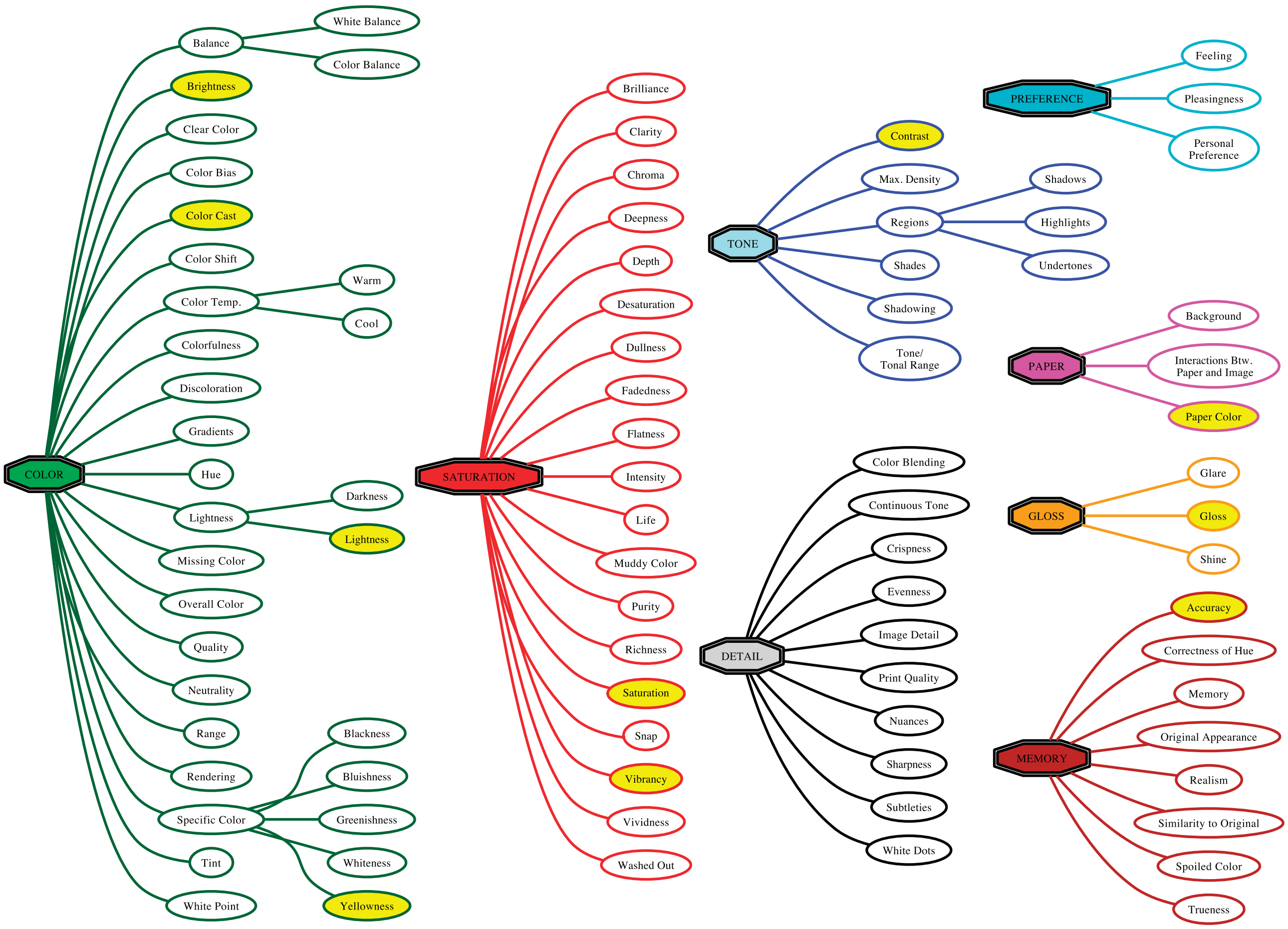


Figure 57. Distribution of words used by participants to describe their Color Rendering Quality decision reasoning

Figure 57 shows the breadth of vocabulary used by observers. This flowchart is most useful for analyzing the variety of word usage. However, the frequency of word usage across all observers for each experiment was still the most useful analytical tool. The following sections discuss the word frequency results for the four experiments.

Image Quality

The Image Quality Experiment was the first of the four experiments conducted by observers. Thus, this was the observers' first encounter with the books and the first time in which they were asked to form impressions of the images and papers. Observers were allowed to rank the images using any criteria they felt were useful in aiding the decision-making process. Thus, it was expected that observer responses would contain words pertaining to both color rendering quality, surface appearance quality, and general preference. This proved to be the case. Figure 58 shows the distribution of words used by participants in their statements following the Image Quality Experiment. Only those words mentioned by at least two participants are shown. The burgundy bars indicate the word was most relevant to surface appearance quality, the gold bars indicate the word was most relevant to color rendering quality, and the blue bar indicates the word was most relevant to general quality.

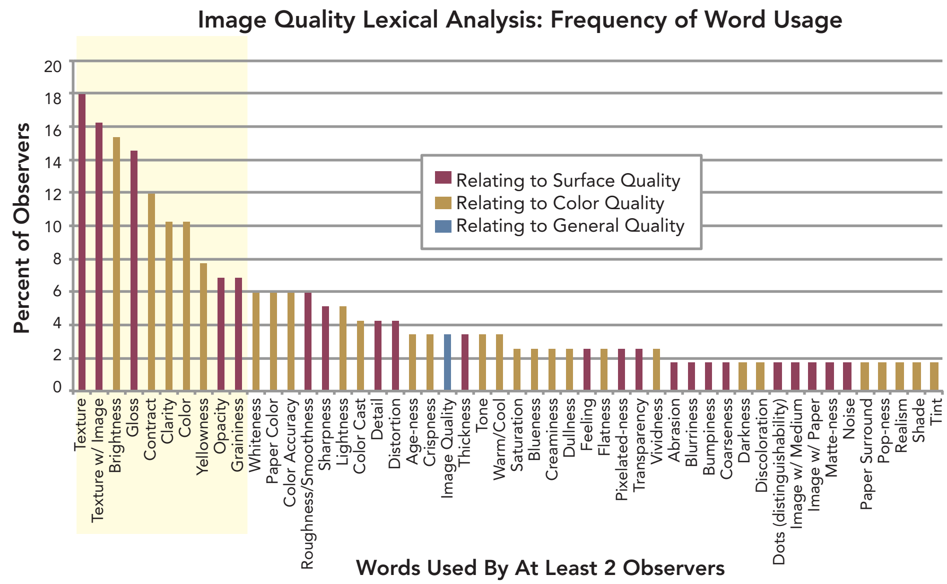


Figure 58. Frequency of word usage collected from observers following the Image Quality psychophysical experiment

As predicted, the number of words relating to Surface Appearance Quality and to Color Rendering Quality were similar: 22 words were related to Surface Appearance Quality and 27 words were related to Color Rendering Quality. Only one word was related to general quality. The bars within the region shaded in yellow (see Figure 58) were the top ten most frequently used words. Within the top ten, five words were related to surface appearance quality and five were related to color rendering quality.

Color Rendering Quality

The Color Rendering Quality experiment was conducted second. Observers were instructed to base their ranking only on factors related to Color Rendering Quality—they could not use factors relating to texture and ink formation. Gloss affects color as much as it does the Surface Appearance Quality; thus, gloss may be considered as applying to both experiments. The frequency plot for words mentioned by observers following the Color Rendering Quality experiment is shown in Figure 59. Of the 51 words mentioned by at least two observers, only four were not specifically related to Color Rendering Quality. Of the four, gloss could be categorized in either case. No words were mentioned relating to general quality.

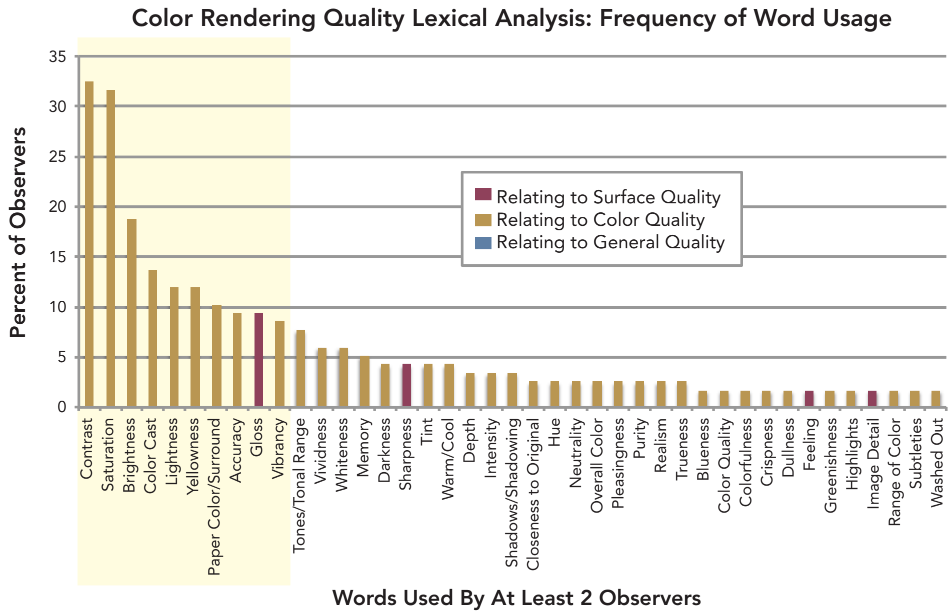


Figure 59. Frequency of word usage collected from observers following the Color Rendering Quality psychophysical experiment

Of the top ten most frequently mentioned words, shown in the yellow region in Figure 59, only gloss was not specifically related to color. This suggests that the observers generally followed the instructions directing them to only consider factors relating to color rendering quality while conducting the experiment.

Surface Appearance Quality

The Surface Appearance Quality experiment was conducted third. Observers were instructed to base their ranking only on factors related to surface appearance quality. Figure 60 shows the results of the word frequency analysis. The number of different words mentioned was the fewest of the first three experiments. Of the 28 words mentioned by at least two people, five did not specifically pertain to surface appearance quality. Two of those five pertained to general quality and three to color quality. Within the top ten most frequently used words, only 'preference,' the tenth most frequently used word, was not specifically related to surface appearance quality. Similar to the results from the Color Rendering Quality experiment, it appears that observers followed the instructions directing them to focus only on factors relating to Surface Appearance Quality, with a few exceptions.

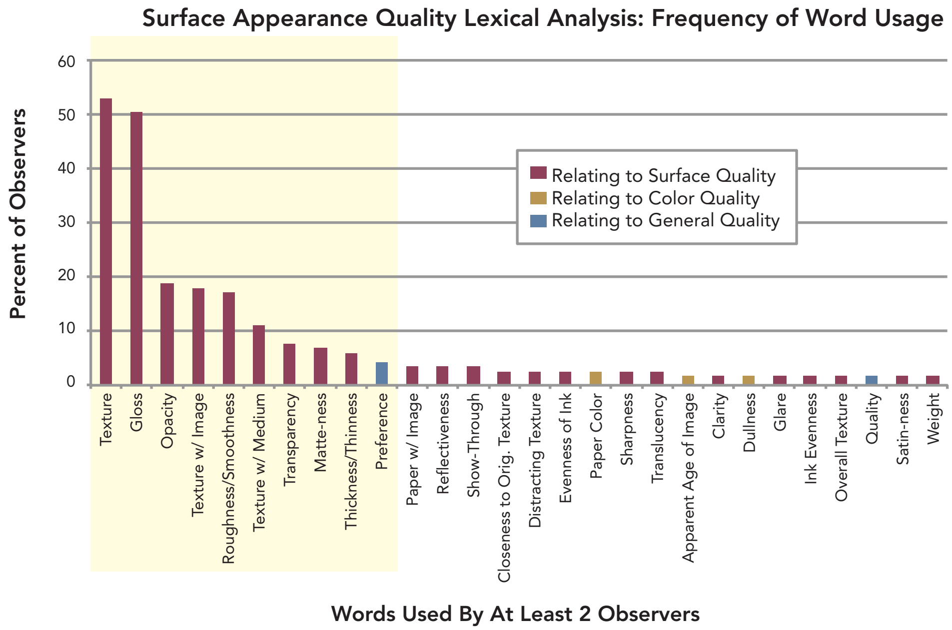


Figure 60. Frequency of word usage collected from observers following the Surface Appearance Quality psychophysical experiment

Customer Preference

The Customer Preference experiment was conducted last. Observers were allowed to look at any or all of the four images in the sample books and were instructed to present the experimenter with their three most preferred books in the order in which they would be inclined to purchase them from a museum gift shop. Figure 61 shows the results of the word frequency analysis for observers' responses to this experiment. The words or ideas used were well distributed between those relating to Surface Appearance, Color Rendering, and general quality. Of the 46 words mentioned at by at least two participants, 18 were related to Surface Appearance, 15 were related to Color Rendering, and 13 were related to general quality. In the three previous experiments at most only two words were related to general quality. However, this is not surprising because observers were instructed to ignore their personal preferences for the first three experiments.

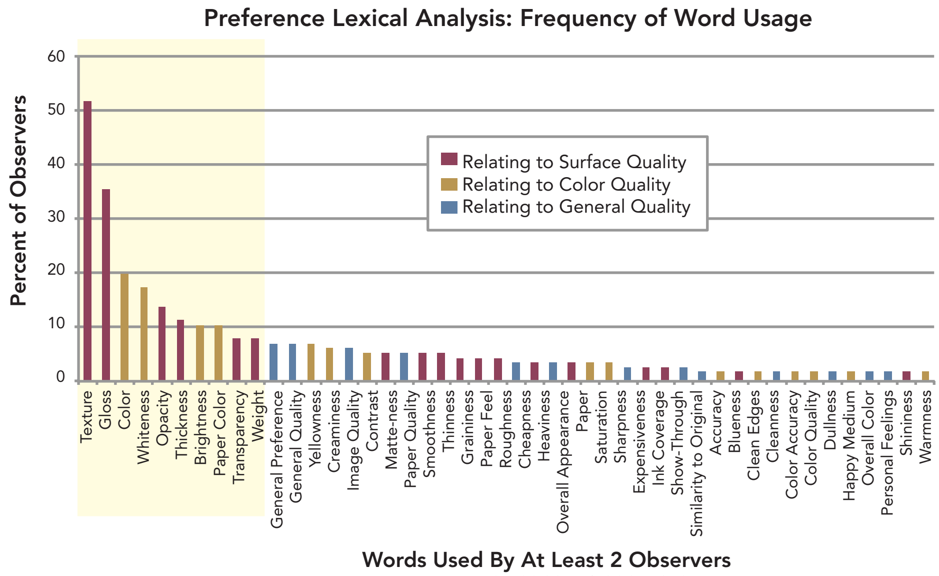


Figure 61. Frequency of word usage collected from observers following the Customer Preference psychophysical experiment

General Analysis

Paper selection is a multi-faceted problem. Many people were interviewed in the first stage of this research. While they represented different fields, they all, with the exception of the paper manufacturer, relied on their knowledge of the art being reproduced and their business knowledge to select paper. Printers, publishers, designers, and curators worked together to choose the best paper for a book without the need for a scientifically and statistically viable experiment. They all understood that paper was rarely selected to produce a facsimile and that art of different media were often printed in the same book. Thus, the main goal for most interviewees was to produce pieces with reasonably accurate reproductions at the best possible price while maintaining readability.

The psychophysical experiment attempted to predict the subconscious reasoning of observers using common print and paper quality metrics. The designed experiment included only coolness, PST, roughness, and gloss as PIPs. The expanded experiment added line raggedness and 40% print mottle as PIPs. The psychophysical experiment provided both specific results and general commentary on the practice of conducting psychophysical experiments. Specifically, a model was developed, in accordance with the Image Quality Circle, to predict the paper properties that would maximize the potential that a paper would be selected for a book. This model is shown in Table 20. Coolness, roughness, and gloss were the significant PIPs for both the designed and expanded experiment models. Those PIPs, weighted differently, were used in both the Color Rendering Quality and Surface Appearance Quality Visual Algorithms. Color Rendering Quality was chosen as the single customer perception in the Image Quality Model. If someone asks which paper is most likely to produce an image with high image quality, then one can say that choosing a paper with low gloss, high coolness and low roughness maximized that chance. However, these data are based upon a small number of papers—though representative of many—printed on two digital presses. These data are significant and come from a large population of observers.

The most significant contribution of this research may be the exploration of statistical design used in conjunction with the Image Quality Circle to conduct a psychophysical image quality experiment. Most documentation does not cite an exploration of the Image Quality Circle to the extent described in this thesis. As a first attempt at conducting this type of experiment, several important results were obtained.

First, there are many confounding factors in paper selection. This experiment limited PIPs to those that are commonly used for print and paper quality. These metrics were never designed to be linearly independent, but only to serve as quality control metrics for the printing and paper industries. Thus, the likelihood that relationships existed between the measured factors was high. It was previously mentioned how PST was confounded with basis weight. This was originally expected and accounted for. However, roughness and gloss were also confounded, although to a lesser extent. The four paper selection parameters—along with basis weight—could all be estimated visually and enhanced the efficiency of paper selection. An alternative method would have been to acquire a large number of papers, physically measure them using a variety of metrics, and select those papers most fitting the statistical design. This method would have required more time than that available for sample selection, and is not a practical method for those without access to a paper testing facility such as that at the RIT Printing Applications Lab. Unlike PST, roughness, and gloss, coolness was specifically designed to measure psychological response. In retrospect, it would have been best to use metrics in place of PST, roughness, and gloss that were designed to measured psychological response.

The second important result related specifically to the design of the psychophysical experiment. The lexical analysis revealed a disconnect between what observers were allegedly thinking and their ranking behavior. The significant PIPs for the Surface Appearance Quality and Color Rendering Quality experiments were the same, suggesting that observers were influenced by the same criteria when ranking the books in both experiments. For example, observers were influenced by texture during the Color Rendering Quality experiment and were influenced by color during the Surface Appearance Quality experiment. Yet, the lexical analysis results suggested that observers were only focusing on factors relative to the specific experiments based upon the fact that the most frequently cited words corresponded to the instructions given at the beginning of the experiment. It is possible that observers framed their post-experiment comments around what they expected was a correct answer. However, whether they were completely truthful in their responses cannot be known. The lexical analysis is assumed to be representative of the observer's conscious decision making. Likewise, the actual rank data is assumed to be representative of the observer's subconscious decision making. Ideally, the two are the same, but that did not seem to be the case.

One plausible explanation for this disconnect was bias resulting from experiment order. The Image Quality experiment was originally chosen as the first experiment because it was reasoned that observers would first judge general image quality while becoming familiar with the samples, then break down image quality into color rendering quality and surface appearance quality. However, while judging image quality, observers may have formed impressions about the pieces that remained throughout the experiment. Under those circumstances, it would have been difficult for observers to ignore factors such as texture or color on which they originally based their decision, while not realizing they were doing so. The solution for future experiments would be to randomize the order of the experiments, ensuring that signs of bias would be discounted by the randomization.

Another plausible explanation for the disconnect between the psychophysical and lexical results may be due to the hints provided to observers to guide them in their ranking process (shown in Table 9). There were no specific hint words provided to observers for the Image Quality experiment. However, seven of the eight hint words provided to observers for the Color Rendering Quality experiment were among the top ten most frequently cited words for that experiment. In addition, the eighth hint word not included in the top ten was cited by more than two observers. Likewise, four of the five hint words provided to observers for the Surface Appearance Quality experiment were among the top ten most frequently cited words. The fifth word was also cited more than two times. Observers may have reverted to using these hint words in their post-experiment responses if they could not otherwise think of suitable words to use. Of course, the hint words are common language and this fact may be confounded with the use of the hint words as crutches.

Furthermore, specific hint words were not provided to observers for the Customer Preference experiment. The wide variety of responses provided by observers illustrates the greater freedom in decision making for this task, similar to that given for the Image Quality experiment. As illustrated by the lexical analysis, participants relied more on general preference for the Customer Preference experiment than for the Image Quality experiment. Yet, the two responses were highly correlated: seven of the top twelve words cited for the Image Quality experiment were among the top ten cited for the Customer Preference experiment. These results suggest that judgments of image quality are not far separated from judgments of preference.

Conclusions

The factors contributing to the selection of paper for books of reproduced fine art were studied and analyzed in this research. This was accomplished through the use of targeted interviews with professionals involved in fine art reproduction and psychophysical experiments examining factors relating to observer perceptions of image quality. Several interesting conclusions were reached based upon an analysis of the targeted interviews. Despite the variety of substrates available for digital production, money is still the most heavily weighted factor limiting the potential for books of reproduced fine art. While larger institutions tend to have larger production budgets, they are not as likely to use that budget for high-quality paper when producing a high volume of books. Smaller institutions or independent publishers may have more flexibility for paper selection when producing short-run, special edition books or books requiring a process unique to digital presses. In general, low costs and production efficiency are the most important for large organizations.

With the exception of paper manufacturers and some printers, most people involved in the process of selecting paper—namely curators, designers, and publishers—have little technical knowledge in the areas of printing and paper manufacturing, and tend to focus more on aesthetics. Publishers and curators are often heavily reliant on designers because they are more familiar with the paper types available and they have a better understanding of the aesthetic relationship between subject matter, printing process, and paper choice. However, the variety of papers is vast, and designers often use what they are familiar with unless a printer or a paper manufacturer introduces them to a new brand. They are less likely to seek out a paper with which they are not familiar. Ultimately, though, the party providing the funds has the final say and can get what they want should they demand it, but in general the relationship between printer, designer, curator or publisher is amicable. The success of the final product depends upon cooperation and mutual respect between those involved in production.

Archivability and green processes are also becoming more relevant in book production. Paper manufacturers and publishers often go to greater lengths to ensure that their products are FSC- or SFI-certified. Such certifications are valuable marketing tools, and, while adding cost in the short-run, can add to a manufacturer's or publisher's reputation and bring in clients who otherwise may look elsewhere. Consumers and cultural heritage institutions are also paying more attention to the longevity of books, especially when considering the cost of many high-quality books of reproduced fine art.

The psychophysical experiment, which analyzed the relationship between paper quality metrics, Customer Perceptions, and Customer Image Quality Ratings, revealed that Image Quality was optimized using papers with high coolness, low roughness, and low gloss. Furthermore, it was found that commonly used print quality metrics, such as mottle and line raggedness, were not significant additions to the Visual Algorithms.

Paper selection is a costly component of book production, especially in the Print-on-Demand industry where workflow flexibility is low. Paper is chosen that will provide the maximum customer satisfaction with minimal cost. In such completely automated workflows it is impractical to make frequent changes to paper in the press trays. Those characteristics revealed in this research to maximize image quality and book preference can help guide Print-on-Demand companies in selecting the optimal paper for printing fine art reproductions. While no one paper will satisfy every customer, using a paper with high coolness, low roughness, and low gloss is the best statistical choice and will ensure satisfaction for the majority of customers.

A comparison between the lexical analysis results and psychophysical analysis results revealed the design did not effectively account for observer bias as a result of experimental order or suggestion due to provided hints. However, these errors are not necessarily failures. They are important steps in understanding the dynamics of running a psychophysical image quality experiment.

An alternate Image Quality Model, which bypassed Customer Perceptions and modeled Customer Image Quality Rating directly using Physical Image Parameters, revealed a different set of optimal parameters. While not the preferred method of analysis for adhering strictly to the Image Quality circle, bypassing Customer Perceptions may have avoided the biases discussed previously. The designed Direct-to-Image-Quality model predicted optimal Image Quality with high coolness, low print-show-through, and low roughness (gloss was not significant). This suggests that observers did, in fact, prefer to see less text printed behind the images they were viewing in the samples. Further testing is needed to determine the implications of using the Direct-to-Image-Quality model.

There are many possible projects that could extend the work described in this thesis. The following section will describe some of those projects.

Future Work

There is currently no standard metric for describing paper color as it is used in paper selection. For that reason, the Coolness Estimation Experiment was created. The Coolness Estimation Experiment developed a metric, called Coolness, for describing the perception of coolness for white paper using colorimetry and UV cut measurements. Coolness was described using a variation of CIE C^*. Not only did the experiment successfully provide a metric for describing participants' perception of coolness for the tested papers, but it also revealed what may be an important insight into how people perceive fluorescent samples. Future experiments will explore why UV-cut measurements describe observer estimates better than UV-included measurements, expand the number of samples, and include multidimensional scaling to determine whether other unanticipated factors are the root of the Coolness Estimation results.

The psychophysical image quality experiment illustrated the vulnerability of the data to human error. The lexical analysis revealed a discrepancy between what observers report as their behavior and how they actually behave. Further experiments will attempt to better control observer learning and bias through randomization of experimental order. In addition, it is necessary to better prepare observers before conducting the experiment so that they understand what will be asked of them, rather than provide clues and hints that may bias the results. It may also be wise to conduct experiments specifically to understand the sources of bias in image quality experiments. While not directly rooted in color science, this will provide the groundwork for the design of future experiments.

Addendum

The claim was made in the analysis of the Coolness Estimation Experiment that *ChromaV1* was the metric best suited for modeling observer perceptions of coolness, as described by Normalized Scale Position. The data was reanalyzed following the completion of the original analysis and it was found that *ChromaV1* was not the most parsimonious model. Rather, it was found that CIE b* alone, calculated using UV-excluded measurements, provided results equal to that of *ChromaV1* (see Figure 62).

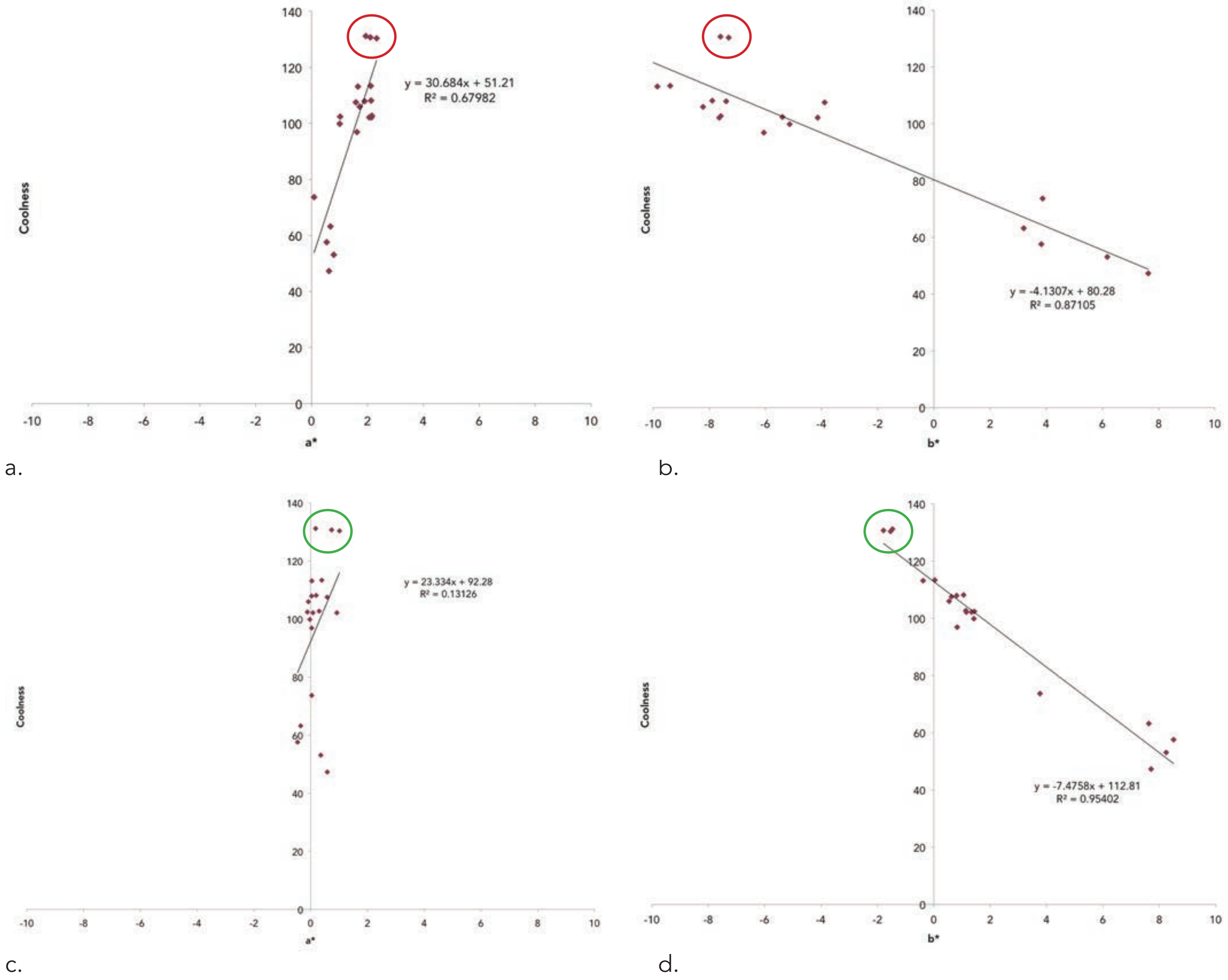


Figure 62. Normalized Scale Position versus (a) a* with UV included, (b) b* with UV included, (c) a* with UV excluded, and (d) b* with UV excluded

Figures 62a and 62b show plots of NSP versus CIE a* and b* calculated using UV-included measurements, and Figures 62c and 62d show plots of NSP versus CIE a* and b* calculated using UV-excluded measurements. Notice the plot in Figure 62b is similar to that in Figure 31a. Likewise, the plot in Figure 62d is similar to that in Figure

31c. In addition, the R^2 values for the latter plots are near identical. Thus, in future examinations of coolness, b^* calculated using UV-excluded measurements will be used instead of ChromaV1. However, the calculates in this research are not compromised, considering that both models predict coolness equally well.

References

Abbott, J. (2003). Publishing trends and libraries. *DESIDOC Bulletin of Information Technology, 23*(1), 7–12.

ANSI/NISO Z39.48. (1992). *ANSI/NISO Z39.48-1992 (R2002) Permanence of Paper for Publications and Documents in Libraries and Archives.* Bethesda, MD: NISO Press.

ASTM D523. (2008). ASTM D523-08 Standard Test Method for Specular Gloss. *Annual Book of ASTM Standards 2009: Section Six, Paints, Related Coatings and Aromatics,* Vol. 06.01.

ASTM D5650. (1997). ASTM D5650-97 Standard Test Method for Resistance to Bending of Paper of Low Bending Stiffness (Taber-Type Tester in 0 to 100 Taber Stiffness Unit Configuration). *Annual Book of ASTM Standards 2009: Section Six, Paints, Related Coatings and Aromatics,* Vol. 06.02.

ASTM D589. (1997). ASTM D589-97 Standard Test Method for Opacity of Paper (15/d geometry, illuminant A/2deg, 89% reflectance backing and paper backing). *Annual Book of ASTM Standards 2010: Section 15, General Products, Chemical Specialties, and End Use Products,* Vol. 15.09.

ASTM D645M. (1997). ASTM D645M-97 Historical Standard. Standard Test Method for Thickness of Paper and Paperboard.

ASTM D646. (1996). ASTM D646-96 Historical Standard. Standard Test Method for Grammage of Paper and Paperboard (Mass Per Unit Area).

ASTM D7163. (2005). ASTM D7163-05 Standard Test Method for Specular Gloss of Printed Matter. *Annual Book of ASTM Standards 2009: Section Six, Paints, Related Coatings and Aromatics,* Vol. 06.02.

ASTM D7305. (2008). ASTM D7305-08a Standard Test Method for Reflection Density of Printed Matter. *Annual Book of ASTM Standards 2009: Section Six, Paints, Related Coatings and Aromatics,* Vol. 06.01.

ASTM D985. (1997). ASTM D985-97 Standard Test Method for Brightness of Pulp, Paper, and Paperboard (Directional Reflectance at 457 nm). *Annual Book of ASTM Standards 2010: Section 15, General Products, Chemical Specialties, and End Use Products,* Vol. 15.09.

ASTM E313. (2005). ASTM E313-05 Standard Practice for Calculating Yellowness and Whiteness Indices from Instrumentally Measured Color Coordinates. *Annual Book of ASTM Standards 2009: Section Six, Paints, Related Coatings and Aromatics,* Vol. 06.01.

ASTM F1944. (1998). ASTM F1944-98 Standard Practice for Determining the Quality of the Test, Line-and Solid-Fill Output Produced by Ink Jet Printers. *Annual Book of ASTM Standards 2010: Section 15, General Products, Chemical Specialties, and End Use Products*, Vol. 15.09.

Axiphos GmbH. (2001, August). *Reviewing the concept of paper brightness.* Retrieved from http://www.axiphos.com/BrightnessReview.pdf

Beretta, G., & Moroney, N. (2008). *Cognitive aspects of color* [Technical report]. Retrieved from http://www.hpl.hp.com/techreports/2008/HPL-2008-109.html

Beretta, G., Moroney, N., & Recker, J. (2009). *Toward robust categorical color perception* [Technical report]. Retrieved from http://www.hpl.hp.com/techreports/2009/HPL-2009-146.html

Besore, C. (Ed.). (1973). *SPIE Handbook of Photographic Science and Engineering* (pp. 1023–1132). New York, NY: John Wiley & Sons.

Bodrogi, P., & Tarczali, T. (2001). Colour memory for various sky, skin, and plant colours: Effect of the image context. *Color Research and Application, 26*(4), 278–289.

Chung, R., & Rees, M. (2006). *A survey of digital and offset print quality issues* (PICRM-2006-04). Rochester, NY: Printing Industry Center at RIT.

CIE, (2004). *Colorimetry technical report* [CIE 15:2004] (3rd ed). Vienna, Austria: Commission Internationale De L'Eclairage.

Coppel, L., Norberg, O., & Lindberg, S. (2010). Paper whiteness and its effect on perceived image quality. *Eighteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications: Vol. 18* (pp. 62–67). Springfield, VA: The Society for Imaging Science & Technology.

de Ridder, H. (1996). Naturalness and image quality: Saturation and lightness variation in color images of natural scenes. *Journal of Imaging Science and Technology, 40*(6), 487–493.

DiSantis, N.E. (2007). *Rub, fold, and abrasion resistance testing of digitally printed documents* (Master's thesis). Retrieved from RIT Digital Media Library. (http://hdl.handle.net/1850/4489)

Engeldrum, P. (1999). Image quality modeling: Where are we? *PICS 1999: Image Processing, Image Quality, Image Capture, Systems Conference: Vol. 2* (pp. 251–255). Springfield, VA: The Society for Imaging Science & Technology.

Engeldrum, P. (2001). *Psychometric scaling: A toolkit for imaging systems development.* Imcotek Press, Winchester, MA.

Engeldrum, P. (2002). Extending image quality models. *PICS 2002: Image Processing, Image Quality, Image Capture Systems Conference: Vol. 5* (pp. 65–69). Springfield, VA: The Society for Imaging Science & Technology.

Engeldrum, P. (2004a). A short image quality taxonomy. *Journal of Imaging Science and*

Technology, 48(2), 160–165.

Engeldrum, P. (2004b). A theory of image quality: The image quality circle. *Journal of Imaging Science and Technology, 48*(5), 447–457.

Evans, M., & LeMaire, P. (2005). *An investigation into papers for digital printing* (PICRM-2005-06). Rochester, NY: Printing Industry Center at RIT.

Fairchild, M. (2006). *Color appearance models.* Hoboken, NJ: John Wiley & Sons.

Farnand, S. (2008). *Minding the gap: Evaluating the image quality of digital print technologies relative to traditional offset lithography* (PICRM-2008-08). Rochester, NY: Printing Industry Center at RIT.

Farnand, S. (2009). *Further investigation into the image quality differences between digital print technologies and traditional offset lithography* (PICRM-2009-04). Rochester, NY: Printing Industry Center at RIT.

Farnand, S., Frey, F., & Anderson, E. (2009). Benchmarking art imaging interchange cycles: Imaging quality experimentation. In D. Smith, Green-Armytage, P., Pope, M.A., & Harkness, N. (Eds.), *Proceedings of the 11th Congress of the International Colour Association* (AIC 2009). CD. Sydney: Colour Society of Australia.

Farnand, S., Jiang, J., & Frey, F. (2010). Comparing hardcopy and softcopy results in the study of the impact of workflow on perceived reproduction quality of fine art images. In S.P. Farnand & Gaykema, F. (Eds.), *Image Quality and System Performance VIII, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 7867,* 7867 05. Bellingham, WA: SPIE.

Fraser, B., Murphy, C., & Bunting, F. (2005). *Real world color management* (2nd ed.). Berkeley, CA: Peachpit Press.

Fredembach, C., Wang, J., & Woolfe, G. (2010). Saliency, visual attention and image quality. *Eighteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications: Vol. 18* (pp. 128–133). Springfield, VA: The Society for Imaging Science & Technology.

Frey, F., Christensen, H., & DiSantis, N. (2006). *Permanence of toner on paper—Based on the life cycle of documents* (PICRM-2006-05). Rochester, NY: Printing Industry Center at RIT.

Hunter, P. (2002). Does calculation of the 95th percentile of microbiological results offer any advantage over percentage exceedence in determining compliance with bathing water quality standards? *Letters in Applied Microbiology, 34*, 283–286.

ISO 2470-2. (2008). ISO 2470-2:2008 Paper, Board and Pulp - measurement of Diffuse Blue Reflectance Factor - part 2: Outdoor Daylight Conditions (d65 Brightness).

ISO 2471. (2008). ISO 2471:2008 Paper and board - Determination of Opacity (Paper

Backing) - diffuse Reflectance Method.

ISO 2813. (1994). ISO 2813:1994 Paints and Varnishes - determination of Specular Gloss of Non-Metallic Paint Films at 20 degrees, 60 degrees and 85 degrees.

ISO 536. (1995). ISO 536:1995 Paper and Board - Determination of Grammage.

ISO 8791-4. (2007). ISO 8791-4:2007 Paper and board - Determination of Roughness/ Smoothness (Air Leak Methods) - Part 4: Print-Surf Method.

ISO/IEC19799. (2007). Information technology method of measuring gloss uniformity on printed pages.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506.

Kelly, K., & Judd, D. (1976). *Color: Universal language and dictionary of names* (National Bureau of Standards Special Publication 440). Washington, DC: U.S. Department of Commerce, National Bureau of Standards.

Kipphan, H. (Ed.). (2001). *Handbook of print media.* New York, NY: Springer.

Kruskal, J., & Wish, M. (Eds.). (1976). *Multidimensional scaling (Quantitative applications in social sciences).* London: Sage Publications.

McCrady, E. (1998). The ANSI/NISO Z39.48 standard and other standards. In E. McCrady (Ed.), *North American permanent papers* (3rd ed.). Retrieved from http://cool.conservation-us.org/byorg/abbey/napp/std.html

Micro Format, Inc. (2010). *Eliminate the confusion: Paper weight and size.* Retrieved from http://www.paper-paper.com/weight.html

Mojsilovic, A. (2002). A method for color naming and description of color composition in images. *Proceedings of the 2002 International Conference on Image Processing, Vol. 2* (pp. 789-792). Piscataway, NJ: IEEE Signal Processing Society.

Mojsilovic, A., Kovacevic, J., Hu, J., Safranek, R., & Ganapathy, S. (2000a). Matching and retrieval based on the vocabulary and grammar of color patterns. *IEEE Transactions on Image Processing, 9*(1), 38–54.

Mojsilovic, A., Kovacevic, J., Kall, D., Safranek, R., & Ganapathy, S. (2000b). The vocabulary and grammar of color patterns. *IEEE Transactions on Image Processing, 9*(3), 417–431.

Mojsilovic, A., & Rogowitz, B.E. (2001). A psychophysical approach to modeling image semantics. In B.E. Rogowitz, Watson, T.J., & Pappas, T.N. (Eds.), *Human Vision and Electronic Imaging VI*, *Vol. 4299* (pp. 470–477). Bellingham, WA: SPIE.

Montag, E., & Kasahara, H. (2001). Multidimensional analysis reveals importance of color for image quality. *Ninth Color Imaging Conference: Color Science and Engineering: Systems, Technologies, and Applications, Vol. 9* (pp. 17–21). Springfield, VA: The Society for Imaging Science & Technology.

Montgomery, D.C., Peck, E.A., & Vining, D.G. (2006). *Introduction to linear regression analysis* (4th ed.). Hoboken, NJ: Wiley-Interscience.

Moroney, N. (2003). Unconstrained web-based color naming experiment. In R. Eschbach & Marcu, G.G. (Eds.), *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VIII, Vol. 5008* (pp. 36–46). Bellingham, WA: SPIE.

Moroney, N. (2009). *Thousands of on-line observers is just the beginning* [Technical report]. Retrieved from http://www.hpl.hp.com/techreports/2009/HPL200959.html

Moroney, N., & Beretta, G. (2009). *Nominal scaling of print substrates* [Technical report]. Retrieved from http://www.hpl.hp.com/techreports/2009/HPL2009322.html

Naudeau, L. (2002). Office copying & printing processes. In L. Nadeau, *Guide to the identification of prints and photographs: Featuring a chronological history of reproduction technologies* (Unpublished book). Retrieved from http:// cool.conservation-us.org/byauth/nadeau/copyingprocesses.pdf

Notable Names Database. (2010). *Chester Carlson.* Retrieved from http://www.nndb.com/people/138/000165640/

O'Neil, E. (2009). *Mother, daughter: Posing as ourselves.* Rochester, NY: RIT Cary Graphic Arts Press.

Pellow, B., Frey, F., & Sorce, P. (2002). *Digital color—Where is the market?* (PICRM-2002-02). Rochester, NY: Printing Industry Center at RIT.

Phillips, J., Burns, P., Fredericks, E., & Rosen, M. (2008). Image quality and cost assessment of print-on-demand books from webbased vendors offering one-off printing. *NIP24: International Conference on Digital Printing Technologies and Digital Fabrication 2008, Vol. 24* (pp. 450–454). Springfield, VA: The Society for Imaging Science & Technology.

Provatas, N., Cassidy, A., & Inoue, M. (2004). Effect of filler distribution and caliper variations on toner transfer in electrophotographic printing. *NIP20: International Conference on Digital Printing Technologies, Vol. 20* (pp. 958–963). Springfield, VA: The Society for Imaging Science & Technology.

Quality Engineering Associates, Inc. (2009). IASLab (Version 2.12.4.0) [Computer software]. Burlington, MA: QEA.

Ramachandran, V., & Blakeslee, S. (1998). *Phantoms in the brain.* New York, NY: Harper Perennial.

Rao, A., & Lohse, G. (1996). Toward a texture naming system: Identifying relevant dimension of texture. *Vision Research, 36*(11), 1649–1669.

Rong, X. (2009). Quality comparison of HP Indigo to offset lithography. *NIP25: International Conference on Digital Printing Technologies and Digital Fabrication 2009, Vol. 25* (pp. 478–481). Springfield, VA: The Society for Imaging Science & Technology.

Searl, M. (Ed.). (2006). *Seeing America: Painting and sculpture from the collection of the Memorial Art Gallery at the University of Rochester.* Rochester, NY: University of Rochester Press.

Siple, P., & Springer, R. (1983). Memory and preference for the colors of objects. *Perception and Psychophysics, 34*(4), 363–370.

TAPPI T410. (2003). TAPPI T410 om-98, Grammage of Paper and Paperboard (Weight per Unit Area).

TAPPI T411. (2003). TAPPI T411 om-97, Thickness (caliper) of Paper, Paperboard, and Combined Board.

TAPPI T425. (2003). TAPPI T425 om-01, Opacity of paper (15/d geometry, illuminant A/2deg, 89% reflectance backing and paper backing).

TAPPI T452. (2003). TAPPI T452 om-98, Brightness of Pulp, Paper, and Paperboard (directional reflectance at 457 nm).

TAPPI T480. (2003). TAPPI T480 om-05, Specular Gloss of Paper and Paperboard at 75 Degrees.

TAPPI T489. (2003). TAPPI T489 om-99, Bending Resistance (stiffness) of paper and paperboard (Taber-type tester in basic configuration).

TAPPI T525. (2003). TAPPI T525 om-92, Diffuse Brightness of Pulp (d/0 deg).

TAPPI T538. (2003). TAPPI T538 om-08, Roughness of Paper and Paperboard (Sheffield Method).

TAPPI T555. (2003). TAPPI T555 om-09, Roughness of Paper and Paperboard (Print Surf Method).

TAPPI T562. (2003). TAPPI T562 pm-96, CIE whiteness and tint of paper and paperboard (using 45/0 directional illumination and normal viewing conditions.

TAPPI T653. (2003). TAPPI T653 om-98, Specular Gloss of Paper and Paperboard at 20 Degrees.

Vogl, H. (2008). *A survey of digital press manufacturers: Critical paper requirements* (PICRM-2006-04). Rochester, NY: Printing Industry Center at RIT.

Waite, J. (2003). Conventional offset lithography and direct to digital printing devices: How do they compete on print characteristics? *IGAEA Visual Communications Journal 2003*, 1–15.

White, D. (2007). *Identification of trends in papers for digital printing* [Technical report]. Atlanta, GA: Center for Paper Business and Industry Studies.

Wolfe, J.M., Kluender, K.R., Levi, D.M., Bartoshuk, L.M., Herz, R.S., Klatzky, R.L., Lederman, S.J., & Merfeld, D.M. (2008). *Sensation and perception* (2nd ed.). Sunderland, MA: Sinauer Associates.

Woolfe, G. (2007). Making color adjustment accessible to non-experts through the use of language. *Fifteenth Color Imaging Conference: Color Science and Engineering: Systems, Technologies, and Applications, Vol. 15* (pp. 3–6). Springfield, VA: The Society for Imaging Science & Technology.

X-Rite. (2011). *DTP70 autoscan spectrophotometer* (discontinued). http://xrite.co.uk/

Xerox Corporation. (1999). *The story of xerography.* Retrieved from http://www.xerox.com/downloads/usa/en/s/Storyofxerography.pdf-

Xu, R., & Kellogg, H. (2007). Print quality of dry-toner color electrophotography for production printing and comparison to offset printing. *NIP23: International Conference on Digital Printing Technologies and Digital Fabrication 2007, Vol. 23* (pp. 378–381). Springfield, VA: The Society for Imaging Science & Technology.

Zeki, S. (1999). *Inner vision: An exploration of art and the brain.* New York, NY: Oxford University Press.

R·I·T

Printing Industry Center

An Alfred P. Sloan Foundation Center

Rochester Institute of Technology
College of Imaging Arts and Sciences
55 Lomb Memorial Drive
Rochester, NY 14623
Phone: (585) 475-2733
http://print.rit.edu