Evaluating the Impact of Printing, Color Reproduction, and Screening Technologies on Image Preference

Mark Samworth[1] and Robert Eller[2]

Keywords: flexography, lithography, expanded gamut, concentric screens

Abstract

A print buyer's choice of printing, color reproduction, and screening technologies may all have an impact on the perceived quality of the resulting images. This paper presents the results of an experiment that assessed the impact that these factors on image preference. Six samples of the same design were prepared using (1) Flexography and Lithography, (2) CMYK (4-color) and Expanded Gamut (7-color) color separations, and (3) traditional (AM) and enhanced (Concentric for lithography, High Definition for flexo) screens. The samples were evaluated by 53 industry professionals who force ranked them under real world conditions (uncontrolled lighting, unscreened participants, etc.). Rankings of participant preference reflected the participant's subjective assessment of image quality and appeal.

A nonparametric test for statistical significance (the Wilcoxson paired difference test) showed several significant differences in participant preference based on the technologies used to produce the samples. Expanded Gamut (7-color printing) was preferred to CMYK regardless of the printing technology or screening used to produce the samples. Offset using Concentric screens was preferred to Flexo using High Definition screens. On the other hand, when High Definition Flexo was compared to standard Offset using AM screens, participants showed no significant difference in image preference.

Introduction

Today, print buyers are presented with a wide variety of options for enhancing image quality. Printing processes, premium screens, and color reproduction strategies are all promoted for their image enhancing capabilities. This paper presents the results of an experiment designed to assess the relative importance of these factors on perceived image quality as judged by a panel of industry professionals.

[1]Esko; [2]Rochester Institute of Technology

Sample Preparation

The samples used in the experiments are summarized in Table 1 below.

Sample	Sample ID	Color Strategy	Screen	Print Technology
1	1020	4-color	200 lpi - High Def	Flexography
2	1024	4-color	200 lpi - AM	Offset Litho
3	1022	4-color	200 lpi - Concentric	Offset Litho
4	1021	7-color	200 lpi - High Def	Flexography
5	1025	7-color	200 lpi - Concentric	Offset Litho
6	1023	7-color	250 lpi - Concentric	Offset Litho

Table 1. *Samples used in the experiments*

All samples were printed using process colors only. Two color separation strategies were used: standard 4-color (CMYK) separations and 7-color Expanded Gamut (CMYKOGV) separations. Two printing technologies were compared in the experiment: Flexography and sheetfed Offset Lithography. For each printing technology, a premium screen was selected to optimize it. For Flexography, a 200 line per inch (lpi) High Definition screen was used. High Definition screens take advantage of 4,000 pixel per inch (ppi) imaging to reproduce up to 400 gray levels at 200 lpi. Offset plates were imaged at 200 and 250 lpi using a Concentric screen. Concentric screening improves image sharpness by dividing individual dots into fine concentric rings. In addition, offset plates were imaged at 200 lpi using a standard Amplitude Modulated (AM) screen.

The test image consisted of the four seed packet designs shown in Figure 1. The test image was color separated using Esko's Color Engine and screened using Esko's AM, High Definition, and Concentric screens. Lithographic samples were printing at McKella 280 in Pennsauken, NJ. Flexographic samples were printed at Clemson University in Clemson, SC.

Figure 1. *Test image*

Experimental Procedures

The experiment was designed to collect preferences from a large sample of industry professionals. To encourage participation, the experimental procedure was optimized for speed and ease of completion. This entailed a number of compromises: 1) participants were not tested to ensure they had normal color vision, 2) the experiment was conducted using ambient (uncontrolled) lighting, and 3) participants were only required to force rank test samples from most preferred (1st place) to least preferred (6th place). The time required for a participant to complete this protocol was less than five minutes.

Once data was collected, the ranks assigned to different treatments were analyzed to determine if the observed differences were statistically significant. A chi square test for goodness of fit confirmed that the data collected was not normally distributed, so a nonparametric test was required. The Wilcoxson Signed Rank test was chosen to analyze the data collected.

Wilcoxson Signed Rank Analysis - Example						
Observer	Sample 1	Sample 2	Sample 1 – Sample 2		Rank	Sign * Rank
			Sign	Abs Value		
A	6	3	+1	3	3.5	3.5
B	5	3	+1	2	1.5	1.5
C	6	2	+1	4	5	5
D	2	5	-1	3	3.5	-3.5
E	6	4	+1	2	1.5	1.5
W^+	Test Statistic W^+ = Sum of Positive Ranks					11.5

Table 2. *Wilcoxson Signed Rank Test – Analytical Procedure*

The Wilcoxson Signed Rank test compares one pair of samples (each representing a unique combination of experimental treatments) at a time. Table 2 shows how the test statistic (W+) is calculated for a small set of sample data. For each pair of samples (e.g., Sample 1 vs Sample 2), the analytical procedure can be summarized as follows:

1. Subtract the scores assigned to one treatment from the scores assigned to the other treatment. In Table 2, Participant A assigned a score of 6 (6th place) to Sample 1 and a score of 3 (3rd place) to Sample 2. The difference in scores is +3 (6-3). Record the sign of the difference (+1) and its absolute value (3) in separate columns.

2. Rank the absolute differences for all participants from smallest to largest. In the event of a tie, average the ranks of the tied differences, and assign this average to each of the tied differences. In Table 2, the smallest difference is 2. This difference would normally be assigned a rank of 1, however, there is a tie. Observers B and E both have a difference of 2. According to our rule for ties, both B and E are assigned the average of the ranks which would otherwise be assigned to the smallest two differences (i.e. ranks 1 and 2). Thus, B and E are both assigned a rank of 1.5 ((1+2)/2). A and D tie for the 3rd and 4th smallest differences, so both are assigned the average of ranks 3 and 4 (i.e. 3.5). Finally, E is unique in holding the 5th smallest difference, and is assigned a rank of 5.

3. Multiply the ranks of the differences by the signs calculated in Step 1 to create the signed rank metric.

4. Calculate the test statistic (W+). The test statistic is the sum of the positive signed ranks. In Table 2, W+ = 3.5 + 1.5 + 5 +3.5 = 11.5.

5. Assess statistical significance. The null hypothesis is that there is no difference between the samples chosen. If there is no difference between Sample 1 and Sample 2, participants will want to award the same rank to both samples. Since ties are not allowed in a forced ranking, Sample 1 will receive a higher rank about half of the time, and a lower rank the other half. In this case, about half of the ranks will be positively signed, and the expected value of W+ (E(W+)) will be approximately half the sum of ranks. Thus, we accept the null hypothesis if W+ is within a two sided confidence interval centered on E(W+). If it is outside this interval (significantly larger or smaller than E(W+)), we reject the null hypothesis and accept the assertion that Sample 1 is different from Sample 2.

For n>20 participants, the distribution of W+ is approximately normal with mean equal to half of the sum of the ranks, and variance equal to one quarter the sum of the squares of the ranks. Since n=53 for this experiment, W+ was tested for significance using the normal approximation just described. A significance level of 95% was chosen for the experiment. Montgomery and Runger (2007) provide a detailed description of the Wilcoxson Signed Rank test.

Results

Experiment 1 was successful in accomplishing its primary objective: attracting a large number of industry professionals who force ranked the test images. Table 5 summarizes the results of Experiment 1 without an assessment of statistical significance.

	Sample					
	1	2	3	4	5	6
Color Strategy	4C	4C	4C	7C	7C	7C
Screen	200HD	200AM	200Con	200HD	200Con	250Con
Print Technology	Flexo	Litho	Litho	Flexo	Litho	Litho
Mean Score	5.0	4.5	3.9	3.1	2.4	2.1
Interpretation	Least Preferred ← ………….…….. → Most Preferred					
Range of Scores	1 - 6	1 - 6	1 - 6	1 - 6	1 - 6	1 – 5

Table 3. *Descriptive Statistics*

As Table 5 shows, there was a definite progression from least preferred to most preferred sample in terms of the mean scores awarded by the participants. On the other hand, the range of scores awarded by the participants was huge. With the exception of Sample 6, every sample was awarded the full range of scores from 1 (best) to 6 (worst). This begs the obvious question, "Are any of these differences statistically significant?".

Wilcoxson Signed Rank Test Results at 95% Confidence Level						
Samples Compared		Sample1	Sample 2	Sample 3	Sample 4	Sample 5
		5.0	4.5	3.9	3.1	2.4
Sample 6	2.1	S6 Pref	S6 Pref	S6 Pref	S6 Pref	Not Sig
Sample 5	2.4	S5 Pref	S5 Pref	S5 Pref	Not Sig	
Sample 4	3.1	S4 Pref	S4 Pref	S4 Pref		
Sample 3	3.9	S3 Pref	S3 Pref			
Sample 2	4.5	Not Sig				

Table 4. *Significant differences in participant image preference*

Table 4 answers this question. This table summarizes the results of applying the Wilcoxson Signed Rank test to all possible pairs of samples. In each cell, either a significant preference is shown or the difference between samples was not found to be statistically significant. In all cases, significance is judged based on a 95% confidence that the results of the experiment are not due to chance. For example, the upper left cell compares Sample 6 to Sample 1 and concludes that Sample 6 is preferred with 95% confidence. On the other hand, the upper right cell compares Sample 6 to Sample 5, an concludes that the difference is not statistically significant.

Twelve of the fifteen possible comparisons demonstrated a significant difference in participant preference. (Note that part of the table has been grayed out to eliminate meaningless self comparisons and redundant mirrored comparisons.)

The results of the experiment are best understood by contrasting the effects of alternative treatments (e.g. color separation strategies). The remainder of this section presents and discusses the results of the experiment by reorganizing Table 4 to contrast alternative treatments.

Expanded Gamut (7C) vs CMYK (4C)					
Color Separation	Printing Technology	Screen	Expanded Gamut (7C)		
			Offset		Flexo
			250 Conc.	200 Conc.	200 HD
CMYK (4C)	Offset	200 AM	**7C Preferred**	**7C Preferred**	**7C Preferred**
		200 Conc.	**7C Preferred**	**7C Preferred**	**7C Preferred**
	Flexo	200 HD	**7C Preferred**	**7C Preferred**	**7C Preferred**

Table 5. *Expanded Gamut (7-color) vs CMYK (4-color)*

Table 5 presents the results of the first comparison, Expanded Gamut (7-color) printing versus CMYK (4-color) printing. Here the results are unambiguous: Expanded Gamut is preferred to CMYK regardless of the printing technology or screen used to reproduce the image. In the past, some individuals have stated that they prefer the softer look of CMYK images to the bold look of Expanded Gamut images. This preference was reflected in the raw data, however, this was a distinctly a minority opinion. Out of 53 participants, only two demonstrated a consistent preference for CMYK. The overwhelming majority of participants, on the other hand, preferred the colorfulness and vibrancy of Expanded Gamut separations. This preference is clearly reflected in Table 5.

Offset vs Flexo (4-Color)		
Print Technology	Screen	Flexo (4C)
		200 HD
Offset (4C)	200 AM	**Not Significant**
	200 Concentric	**Offset Preferred**

Table 6. *Offset vs Flexo (4-Color)*

It is impossible to separate the visual impact of printing technologies from the screens used to prepare their plates, so these two treatments will be discussed together. Table 6 summarizes the results of comparing 4-color Offset images to 4-color Flexo images. An important negative finding is that there was no significant difference in participant preference between standard Offset (using a 200 lpi AM screen) and 200 lpi High Definition Flexo. In other words, a panel of 53 industry professionals were indifferent between the quality of standard Offset and high end Flexography. Offset, of course, has not been standing still. When 200 lpi High Definition Flexo was compared to 200 lpi Concentrically screened Offset, Offset was preferred.

Flexo vs Offset (7-Color)		
Print Technology	Screen	Flexo (7C)
		200 HD
Offset (7C)	200 Concentric	**Offset Preferred**
	250 Concentric	**Offset Preferred**

Table 7. *Offset vs Flexo (7-Color)*

Table 7 compares high end Flexo to high end Offset when printing 7-color separations. Based on the results shown in Table 6, it comes as no surprise that high end Offset is preferred over high end Flexo.

Offset Screens (4-Color)		
Print Technology	Screen	Offset (4C)
		200 Concentric
Offset (4C)	200 AM	**Not Significant**

Table 8. *4-Color Offset Screens*

Although the use of a premium screens (Concentric screening) was a significant factor in differentiating high end Offset from high end Flexo, the same cannot be said when Offset is compared to Offset. As Table 8 shows, when 4-color offset images prepared using a standard 200 lpi AM screen were compared to images prepared using a 200 lpi Concentric screen, the difference in participant preference was not statistically significant.

Offset Screen Frequency (7-Color)		
Print Technology	Screen	Offset (7C)
		250 Concentric
Offset (7C)	200 Concentric	**Not Significant**

Table 9. *7-Color Offset Screen Frequency*

Finally, 200 lpi Concentric screening was compared to 250 lpi Concentric screening when printing 7-color separations using a sheetfed Offset press. As Table 9 shows, increasing screen frequency did not result in a statistically significant difference in participant preference.

Conclusions

A panel of 53 Graphic Arts professionals:

1. Preferred Expanded Gamut (7-color) prints over CMYK (4-color) prints regardless of the printing or screening technology used to render the images.

2. Preferred Offset images prepared using a premium Offset screen (Concentric screening) over Flexo images prepared using a premium Flexo screen (High Definition screening).

3. Were Indifferent (did not have a statistically significant preference for either image) when a standard Offset image (200 lpi AM screening) was compared to a premium Flexo image (200 lpi High Definition screening).

4. Were Indifferent (did not have a statistically significant preference for either image) when a standard Offset image (200 lpi AM screening) was compared to a premium Offset image (200 lpi Concentric screening).

5. Were Indifferent (did not have a statistically significant preference for either image) when an Offset image prepared at 200 lpi using Concentric screening was compared to an Offset image prepared at 250 lpi using Concentric screening.

Literature Cited

Montgomery, D. C. and Runger, G. C.
2007, Applied Statistics and Probability for Engineers, 4th Edition, John Wiley and Sons, Inc.