Understanding the Role of Print Quality in Perceived Printer Quality

Paul L Jeran and Abbie J. Parker
Hewlett-Packard Company, Boise, Idaho

Steven V.C. Korol
Xerox Corporation, Wilsonville, Oregon

Abstract

As difficult as it may be for the imaging science community to accept, there is more to perceived color printer quality than print quality. For a moment, place the magnifying loupe aside and check the colorimeter at the door – this paper takes a look at the bigger picture. Several *other* attributes influence the likelihood a potential customer will select one printer over another. The following paragraphs discuss a set of attributes that more completely characterize overall printer quality. Attributes considered relate to print quality, cost, speed, and reliability. Printers used in the experiment span the technologies commonly found in the business color printing market: electrophotographic, solid ink, and drop-on-demand ink jet. Psychometric evaluation methods are employed both to quantify the print quality of test samples as well as to develop the relationships between the chosen attributes. Examples are given of how such an understanding of these interrelationships may provide guidance in making trade-offs during the design process. In conclusion, the authors highlight their estimate of the importance of print quality to printer purchase likelihood.

Introduction

In developing complex, technology rich products, focusing on any one performance metric during the development process is, to say the least, not advisable. Color printers sold into the office market are certainly no exception to this rule. Still, as imaging scientists and engineers, we often focus our attention on a very small subset of the attributes that define product quality as perceived by our customers.

While hardcopy appearance is certainly of great importance to the development of a successful printer, it is only one line item from a lengthy list of factors – all of which are important to our customers. For example, device cost, reliability, and cost of consumables can have as much or more of an effect on a successful product than print quality. A printer with market leading print quality but costly consumables may not be as successful as a printer of average print quality but low cost of consumables – point being that in the design of a printer, a *systems approach* must be taken to assure success in the market place.

Understanding the interrelated nature of the key attributes of a printer system provides a tool for making trade-offs in the design process. In order to achieve this understanding, key attributes were first defined and subsequently quantified where necessary. Given these key attributes, a factorial series of psychometric experiments was carried out with the goal of developing a model of the relationships between the attributes. This model was then used to determine the relative contributions of the various attributes to overall printer quality.

Experimental Method

While the number of attributes that may actually influence an individual's decision to purchase a printer is quite large, it is the opinion of the researchers that there are a handful of key attributes that dominate the decision making process. These dominant attributes were selected for consideration in the experimental series: print quality, cost of ownership, effective throughput performance, and reliability.

Of these attributes, print quality is the trickiest to quantify. Device addressability, quoted in dots per inch (dpi), has become a common specification given to represent a level of print quality. In actuality, addressability may have very little bearing on the perceived quality of prints generated by the device. Rather than such specs, psychometric scaling methods were used to rate print samples according perceived print quality. In such psychometric experiments, observers look at actual print samples rather than specifications, yielding a much more accurate picture of the customer's perception of true print quality. Given the care taken to capture true print quality, it followed to give more information than just specifications on the other attributes. Rather than reviewing traditional specifications, observers in the experiment were asked to consider slightly different terms to describe how well a user might actually perceive a printer's throughput performance, cost of ownership, and reliability.

Ranges for each attribute were determined by general assessments of a variety of currently available color printers targeted at the small office to enterprise office workgroup printing environment. Technologies considered included electrophotographic, drop-on-demand ink jet, and solid ink.

Print Quality

Quantifying print quality of the printers used in the experiment consisted of first constructing a print sample document, printing it on several actual devices using various quality and speed settings, and completing a print quality ranking experiment to determine the quality differences between the samples. To an inexperienced reader, such a series of steps may sound fairly trivial; however, to one who has spent any time in the trenches of print quality evaluation, surely the mere mention of the subject simultaneously piques the interest and raises the little hairs on the back of one's neck.

Print Quality Test Sample

Significant time was invested in each of these steps, beginning with the construction of the test print. With the inherent subjectivity of visual evaluations, it is often an excellent idea to keep the number of variables presented to the observer to a minimum. In the *best* case for the experimenter, the observer would have only one variable, or attribute, to rate. In the real world, though, such simplicity is rarely ever the case. Even in the simplest of text documents, several variables may exist, from jagged or blurred edges to background toner, etc. For the target customer of this experiment, several print quality attributes would be of high importance. Attributes such as text quality, color saturation, banding, halftone-induced patterning, and graininess are just some of the factors that play into perception of print quality. To isolate each of these attributes one per print and present several prints at several attribute levels would have proven both intractable and unrealistic. Therefore, a single print was designed to incorporate many of the attributes in a customer-type image. Observers were permitted to evaluate the samples based on any criteria or combination thereof they desired, and statistics were used to guide the experimenters in rating the samples according to print quality based on the responses of the observers.

The print sample was designed to mimic a customer-type print while stressing the print quality capability of the devices used to generate the prints. In order to accomplish these goals, the document turned out to be fairly complex, consisting of mixed text and graphics, with both monochrome and color components. In addition to multi-colored text and line art, pictorial images containing memory colors (flesh, sky, and sand) completed the page. Figure 1 presents the test image. The original was full color, and nominally consisted of roughly 15% per color for each of the process colors and black. Of course, the actual amount of colorant applied to the page by any printer is highly device dependent, so these percentages varied, but all were in this neighborhood. While 15% per color is substantially higher than the generally accepted 3-5% average color coverage page for this market, it was deemed necessary to permit observers to evaluate the prints based on many print quality attributes likely to be encountered in office printing. Further, it was the belief of the experimenters that some or all of the attributes within this type of complex image are likely to be used to evaluate the print quality of a device by real-world customers. Sure enough, notes taken by the observers during the experimentation confirmed that in many cases, several attributes were weighed against one another in order to rate the print quality of the samples.

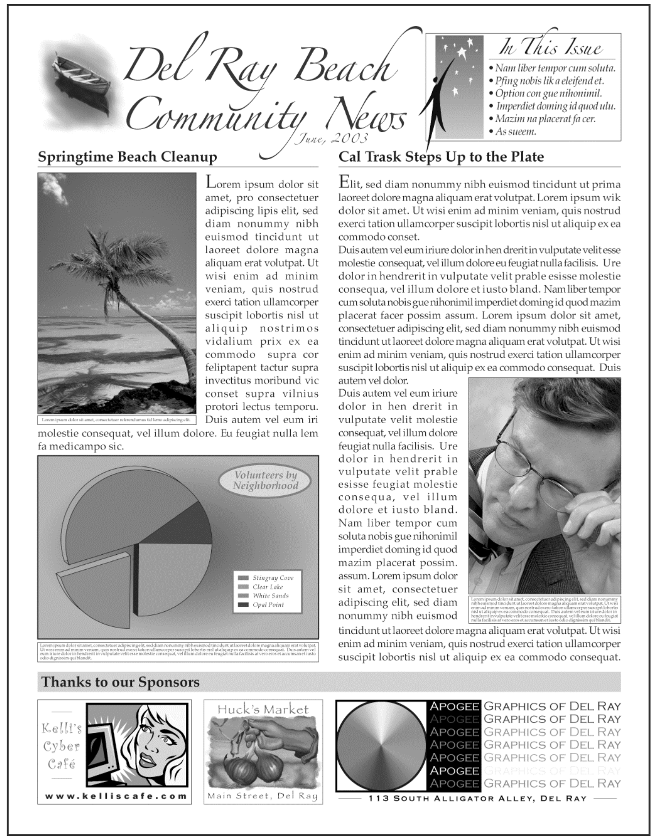

Figure 1. Image used for print quality evaluation. Original image was full color.

Print Quality Scaling

Given the standardized test image, a series of prints was made using several color printers using various output quality and speed settings. All prints were made on commonly available 24lb (90 gsm), 94 brightness Hammermill® LaserPrint™ paper. Resulting prints ranged considerably in print quality. None of the samples were processed in any way to exaggerate their quality – the range of print quality bracketed that found in the current office color printer market. The goal of the print quality evaluation experiment was to yield a small set of quantified, well-differentiated samples to be used in further scaling experimentation with the three aforementioned additional attributes of cost of ownership, effective throughput performance, and reliability.

In order to quantify differences in print quality between the samples, the experimenters employed the well-known technique of paired comparisons combined with Case V of Thurstone's Law of Comparative Judgment (referenced and explained in several texts, including Engeldrum, 2000). This method was chosen primarily because the print samples were both complex and more similar to one another than they were different. When such is the case, that is, when samples are similar enough to be confused often, paired-comparison methods work best, i.e. much better than rank ordering (Bartleson and Grum, 1984).

The experiment consisted of sequentially presenting randomized pairs of samples to each observer and asking for an ordinal judgment between the two. Specifically, the observers, isolated in a room with a viewing easel and D50 lighting, were instructed:

> *You will be presented with a series of print pairs in random order. Your task is to indicate with a check mark on the graduated scales your preference for one print or the other, and to what degree.*

Every possible unique combination of sample pairs was judged by each observer for a total of N = $(n/2)(n\text{-}1)$ observations, where n represents the number of samples in the set. A total of eight samples was chosen as a good compromise between providing sufficient coverage over the range of print quality studied and limiting observer fatigue (Lyne, 1979), for in this method, eight samples translates to 28 print-pair observations for each observer. Each print-pair observation yielded a proportion; it was through analysis of these proportions that the scale values were estimated. 30 observers performed the experiment to assure statistically adequate comprehension of the print differences.

From this data, an interval print quality scale was created. Figure 2 presents these results, with the output of the ranking experiment, Normalized Estimated Ranking, plotted against sample identification. Of course, the data shown is actually one-dimensional, consisting of the intervals between the samples and the order of the samples relative to one another, which was precisely the goal of the experiment. Although beyond the scope of this paper, interesting two-dimensional plots based on the same data would cast the scale values against various physical image parameters, such as optical density or modulation transfer. Such plots would provide insight into the importance of those parameters to perceived print quality – an interesting topic to save for another day.

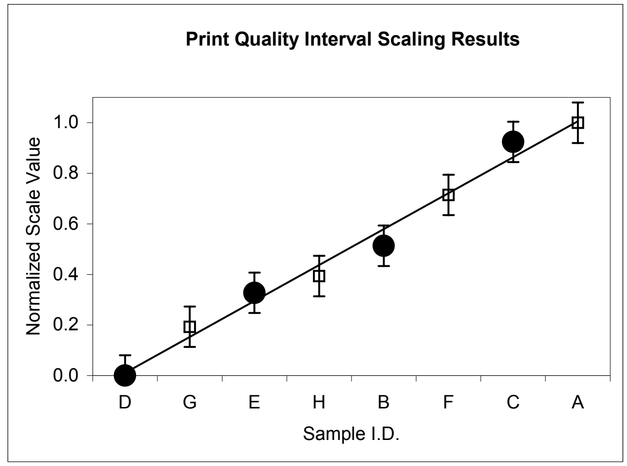

Figure 2. Print Quality Interval Scaling Results from the Paired-Comparison Print Quality Study. Sample "D" is of the highest print quality.

The error bars around the points of Figure 2 represent greater than 90% confidence intervals for the points, and were derived from the data of Figure 3. Torgerson (1958) proposed a convenient method of checking the goodness of fit of the model to the data by running the model in reverse. After deriving the scale value estimates (plotted in Figure 2) from the experimentally obtained proportions, the scale value estimates were then used to calculate model predicted proportions. In the diagnostic plot of Figure 3, the Experimentally Observed Proportions are plotted against the Model Predicted Proportions. The width of the distribution of these points about a line of unity slope reveals the error between the actual observed proportions and those predicted by the model. In fact, with the 28 proportions of the experiment plotted, the width of the distribution represents a 93% confidence interval. The error bars of Figure 2 equal this width.

Points on the graph of Figure 2 marked with black dots represent those samples that were used in subsequent scaling experiments. The error bars reveal that care was taken in their choice to be sure that each was statistically different from its neighbors. The intervals between the samples guaranteed that the perceived print quality level of the chosen samples evenly spanned the print quality range under study.

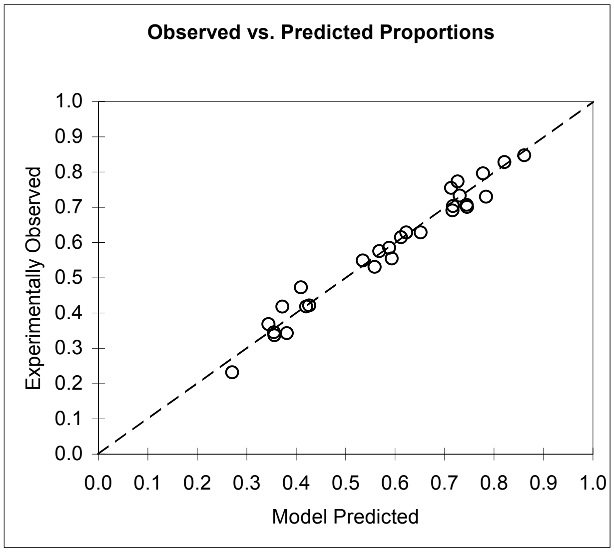

Figure 3. Diagnostic plot of Experimentally Observed vs. Predicted Proportions.

Effective Throughput Performance

Traditional printer throughput performance metrics used in the industry are: continuous printing speed in pages per minute (ppm), time to first page out (FPO), and sometimes processor speed (given in MegaHertz, or Mhz). Observers were asked to consider a printer's *effective throughput performance*. The definition stated that effective throughput performance did not represent the rated engine speed that typically lists on a spec sheet. Rather, the effective throughput performance takes into consideration first page out time, processing efficiencies, software or engine pauses and other common system components that cause a printer to print slower than the rated engine speed.

Effective throughput performance represents what general office users might actually see for printer throughput given a mix of document complexities, applications, job lengths and so on. For example, a printer with a high rated speed, but slow processing capabilities and slow FPO time may actually print a selection of files in about the same amount of total time as a printer with a slow speed, but fast FPO and fast processing capabilities. Even though these two printers have very different performance specifications, an office user might realize the same effective throughput performance on both devices for a given set of print jobs. The three levels tested were 3ppm, 8ppm and 14ppm.

Cost of Ownership

Cost of ownership may be calculated using a long list of factors, from power consumption to labor of an operator. In order to minimize items to consider, the dollar amount listed for cost of ownership in this experiment represented the sum of the printer purchase price, the cost for an extended 3 year warranty, and the cost of consumables (all supplies, except paper) to operate the device over a 3 year time period when printing 2,500 pages per month. Observers were told to assume 50% of the documents printed are color and 50% are monochrome. The three levels tested were $5,000, $8,000 and $11,000.

Reliability

Reliability is perhaps the most difficult of the attributes to define and capture with a single, understandable value. Printer specifications sometimes call out monthly duty cycle to support claims of reliability. Because it is so difficult to quantify reliability using the limited information provided by a manufacturer or for an average user to test, the researchers attempted to give a sense of how reliable a printer is with how often a user will have to interact with it. The experiment referred to this as *intervention frequency*.

The frequency of intervention describes how often a printer requires interaction from a user beyond simply adding more paper into a tray. This number gives a high level indication of how an end user will perceive the reliability of a printer after using it over time. Actions included for the intervention number are clearing a paper jam, changing a supply (toner, ink, etc.), maintenance requirements, or even a service call to fix a problem.

Including many factors into the intervention frequency creates a high level view. This approach may undermine the severity of one issue compared to the simplicity of another issue or, conversely, amplify the importance of a simple interaction compared to a difficult task. In other words, a printer with high capacity cartridges that breaks down frequently may be assigned the same intervention frequency as a printer with very low capacity toner/ink cartridges but that does not have frequent problems. The three scenarios tested were the following: once per week (or, once every 625 pages printed), once every three weeks (or, once every 1,875 pages), and once every six weeks (or, once every 3,750 pages).

Determining the Overall Perception of Printer Quality

In order to comprehend the relative importance of the factors under study to a customer's purchase decision, it was necessary to understand not only the independent effects of each of the variables, but also the effects of their interactions. Studying all the effects of four different print samples together with three levels each of the other variables necessitated a factorial experiment. The experimental series consisted of 36 hypothetical printers, identified by a print sample and information card. All varied in print quality, speed, cost of ownership, and reliability.

Category scaling was used to rate the likelihood of purchase for each of the 36 theoretical printers. In this test, observers were given seven categories and asked to place each printer in the bin corresponding to the likelihood that they would purchase the printer. The categories were:

Very likely to purchase
Likely to purchase
Somewhat likely to purchase
May purchase
Somewhat not likely to purchase
Not likely to purchase
Not very likely to purchase

The charge given to observers was:

> *Today, we are asking that you place yourself in the position of a small to medium business owner who is looking to buy a color printer to be used in a shared workgroup office environment. With any decision, there are tradeoffs, or pro's and con's. You will be looking at several hypothetical color printers. In addition to looking at a print sample for each printer, you are provided with information on printer performance, cost of ownership and reliability for each. The definition of each of these attributes is described below. After examining each of the hypothetical printers, please place it in one of seven categories listed on the table to indicate how likely you are or are not to purchase the printer.*

Additionally, the observers were provided with the definitions for effective throughput performance, cost of ownership, and reliability.

Thirty observers participated in the experiment. At the end of the experiment, a ratio scale of purchase likelihood was created for each of the 36 hypothetical printers. The main effects and interactions thereof were analyzed to better understand the purchase decision. The sorted standardized pareto chart in Figure 4 represents the estimated effect divided by the standard error. The vertical line represents the cutoff for effects that are statistically significant. Attributes with a standardized effect past the line indicate a significant effect on the likelihood to buy. While all attributes play a significant part in the decision to buy, it is interesting to note that the effect of reliability is similar to the effect of the interactions between print quality and cost, and print quality and speed.

An alternate method to determine the significant effects is to plot all of the effects on a normal probability plot. The effects near the line can be explained by normal variation, while those far from the line are considered to be significant effects. From this approach, it is clear that print quality, speed, and cost are the most significant effects, while the importance of reliability to purchase likelihood is much more difficult to distinguish.

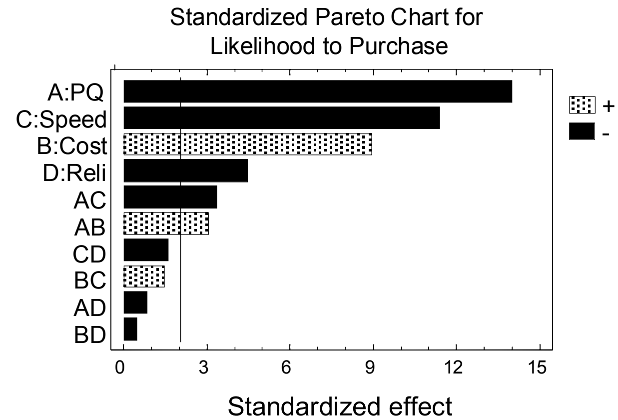

Figure 4. Sorted Standardized Pareto Chart for Purchase Likelihood. Effects extending beyond vertical line are significant.

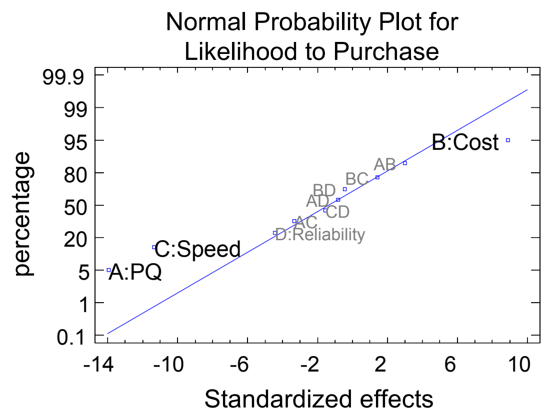

Figure 5. Normal Probability Plot for Purchase Likelihood. Print Quality, Speed, and Cost are primary drivers.

Conclusions

As expected, print quality is indeed an important factor in a color printer purchase decision. In fact, statistically, it is the *most* important factor – which came as a bit of a surprise to the experimenters, who had almost expected to see speed or cost at the top, given the decent print quality of even the lesser quality samples. One possible reason for this finding (beyond simple truth) was that it may have been easier for the observers to relate to the print quality samples than the other attributes, due to the fact that all other factors were presented as numbers on a card, rather than real examples. If the observers actually watched two printers of similar print quality, cost, and reliability printing, one at 3ppm and the other at 14, the conclusion may have been slightly different.

Although reliability plays second fiddle to print quality, speed, and cost, it is unreasonable to conclude that reliability is not an important factor to customers. Rather, for the levels tested in this experiment, and the definition presented, customers perceived the intervention frequency as adequate for this class of printer.

The authors recognize that there are a number of other factors beyond those tested in this study that can make or break a printer's success in the market place. Such is especially evident if an attribute does not meet a minimum threshold requirement. For example, a customer in need of a minimum of a 250-sheet input capacity would never consider a device having only 50-sheet capacity even though it meets or exceeds all other expectations.

It is also interesting to point out that many customers do not actually evaluate print samples when selecting a printer, but rather use specifications as their guide. Such is somewhat true for the other attributes as well, as they were defined in terms much more akin to actual usage. While the researchers are not prepared to comment on the effects of this inconsistency with the actual purchase process, it would seem logical that the conclusions of this research may relate more to the perception of quality by the customer after the sale than prior.

Lastly, care must be taken not to demerit speed and cost, but rather to see them as factors nearly as important as print quality – they were found to have a substantial influence on the likelihood to buy. The results strongly suggest that, for this market segment, cost and speed should be considered nearly equally with print quality.

Acknowledgements

The authors wish to express their sincere appreciation to Dr. Marcos Esterman of Hewlett-Packard for providing expert consultation in the area of DOE.

References

1. Bartleson, C.J., Grum, F., Optical Radiation Measurements, Academic Press, Inc., Orlando, FL, 1984.
2. Box, George Edward, Hunter, Willaim G., et. al., Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building, Wiley, John & Sons, Incorporated, New York, NY, 1978.
3. David, H.A., The Method of Paired Comparisons, 2nd Ed., Oxford University Press, New York, NY, 1988.
4. Engledrum, P.G., Psychometric Scaling, Imcotek Press, Winchester, MA, 2000.
5. Lyne, M.B., Multidimensional Scaling of Print Quality, Tappi Journal, Vol. 62, No. 11, pp. 103-107, 1979.
6. Torgerson, W.S., Theory and Methods of Scaling, Wiley and Sons, New York, NY, 1958.

Biographies

Paul Jeran joined Hewlett-Packard in 1992. During that time he has been involved in the development of new printing technologies, print quality measurement and printer reliability. Currently he is a member of the Advanced Technology Section working to establish print quality metrics for LaserJet printers. He is an active member in ISO/IEC JTC1 SC28 standards committee. Mr. Jeran received his BS and MS degrees in Mechanical Engineering from the Rochester Institute of Technology.

Steve Korol is a principal research and development engineer in the Xerox Office Group, where he manages the Advanced Printhead Development Group. He has worked for Xerox since their acquisition of the Tektronix Color Printing and Imaging Division in 2000. During this time, he has focused on solid ink technology, where he holds a few patents. Steve worked for Tektronix from 1994 until the Xerox acquisition as an imaging scientist, focused on color print quality evaluation and optimization. He has MS and BS degrees in Imaging Science from the Rochester Institute of Technology. Steve is committed to the IS&T, and in addition to teaching tutorials and presenting papers, he contributes to the Technical Leadership Committee and has served various functions on numerous conference committees over the last several years.

Abbie J. Parker began working for Hewlett-Packard in 1998 on contract and became an employee in 2000. During this time she has worked on testing, analysis and technology research for future product development. She is currently a member of the Advanced Technology Section and her main area of focus is color page printer technology. She is an active member in ISO/IEC JTC1 SC28 standards committee. Ms. Parker received her BS degree in Chemical Engineering from the University of Idaho.