Scholars' Mine

Masters Theses

Student Theses and Dissertations

Spring 2010

Quantification of product color preference in a utility function

Hannah L. Turner

Follow this and additional works at: https://scholarsmine.mst.edu/masters_theses

Part of the Manufacturing Commons

Department:

Recommended Citation

Turner, Hannah L., "Quantification of product color preference in a utility function" (2010). *Masters Theses*. 4780.
https://scholarsmine.mst.edu/masters_theses/4780

This thesis is brought to you by Scholars' Mine, a service of the Missouri S&T Library and Learning Resources. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

QUANTIFICATION OF

PRODUCT COLOR PREFERERENCE

IN A UTILITY FUNCTION

by

HANNAH L. TURNER

A THESIS

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN MANUFACTURING ENGINEERING

2010

Approved by

Seth D. Orsborn, Advisor
Katie A. Grantham
Shun Takai

© 2010
Hannah L. Turner
All Rights Reserved

ABSTRACT

Currently, many marketing and engineering tools exist to help a designer optimize quantitative attributes of a product, such as height, weight, volume, or cost. However, these methods cannot effectively take into consideration aesthetic attributes of a product, or any other attributes for which there is no understood functional relationship between the attribute's potential values and the consumer's preference.

This research has begun the work of developing this necessary functional relationship for the aesthetic attribute of color and has created a methodology for further research. To do this, colors were represented by their red, green, and blue light components, and preference information for each of these attributes was gathered by presenting individuals with a small sample of colors, applied to backpacks, in a short choice survey. A utility function was fit to the preference data points using standard regression methods.

The validity of these functions was tested by administering individual-specific follow-up surveys, in which each of the survey questions contained a high, a neutral, and a low utility backpack color, as determined by the utility functions. Individuals chose the high utility color an average of 74% of the time, which is significantly better than random chance. In addition, success rates as high as 87% were achieved in certain instances where greyscale preferences were incorporated into the overall utility function. These results indicate that a large portion of individual preferences were captured by the utility functions, allowing the methodology provided to serve as a foundation for future research.

ACKNOWLEDGMENTS

Thank you to Dr. Seth Orsborn for his patience and guidance as my advisor throughout this research, and to Dr. Katie Grantham and Dr. Shun Takai for investing their time as committee members.

Additional appreciation is given to Rachel Day, Garrett Foster, and Tony Talecki for their assistance with data collection, and to Stephen Mues for his truly invaluable computer programming skills. Finally, thank you to SurveyGizmo for generously offering the use of their high-end online survey software free of charge to students.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Page

1. INTRODUCTION

With the emergence of internet retailers and competition from global suppliers, it is increasingly necessary for companies to design and manufacture products that meet consumers' wants and needs on every level. Tools such as the Quality Functional Deployment [1] have helped engineering designers to translate customer needs into product functionalities, giving designers a means to understand the inherent trade-offs involved in a design (for example, between maximizing the capacity of an aircraft and minimizing its weight) and to develop one or more functionally optimal products. However, these methods cannot effectively take into consideration aesthetic attributes of a product, or any other attribute for which there is no understood mathematical relationship between the attribute's potential values and the consumer's preference. For this reason, aesthetic design decisions are typically left up to creative experts who rely on a combination of design heuristics, current trends, and educated intuition when making decisions about a product's aesthetics [2]. Without any kind of proof to validate these choices, engineers are unlikely to give aesthetic attributes fair consideration when products must be redesigned to reduce costs or increase manufacturability. However, product aesthetics can make up 40 – 90% of a consumer's purchase decision [3], and these aesthetic compromises can create failures out of functionally acceptable designs.

These issues are particularly prevalent when it comes to determining a product's color. While much research has been done on the subject of color preferences, the focus has been almost entirely on determining the universal preference order of colors, and how those preferences change for different genders, cultures, or age groups. That is to say, research in color preferences has sought to understand these preferences for purely academic purposes, separate from any application to the design of actual products and without any intent for optimization.

The need exists, then, for a method that can quantifiably represent consumers' color preferences with respect to measurable color attributes. This can be done using utility functions, where the measurable color attributes are the red, green, and blue light components that combine to create colors in the visible spectrum. Utility is a measure of satisfaction that a consumer has for a given product, and it can be represented as a function of the measured values of each of these attributes. These attribute utility functions can be developed by obtaining preference information from individuals through a simple discrete choice survey. Optimization techniques can then be applied to these equations to generate the product color most preferred by any given individual. In other words, this method collects a limited amount of information from the consumer and in turn provides the designer with a consumer's preference for any color in the visible spectrum, reducing the amount of time, money, and subjectivity involved in the determination of product colors.

2. REVIEW OF LITERATURE

In order to effectively model consumer colors preferences, an introduction to the basic concepts and current research regarding preference modeling and utility theory is useful. Additionally, prior works in the field of color preferences are also discussed in the following sections.

2.1. PREFERENCE MODELING AND UTILITY THEORY

Identifying customer needs and preferences and accurately translating them into a product's features and functionalities is essential to successful product design. Although many valuable methods exist to aid the designer in this part of the process, none of the currently available methods is fully able to incorporate qualitative preferences (such as those for aesthetics or usability) due primarily to their non-numeric nature.

For example, the widely used Quality Functional Deployment, or House of Quality [1] provides a means to translate customer needs to measurable technical requirements which designers can then attempt to maximize, minimize, or target to specific values. In this method, however, customer needs such as "be visually appealing" are difficult, if not impossible, to incorporate into this model without measurable methods for representing factors, such as color or form, that contribute to visual appeal.

The issue of translating and interpreting customer needs is further complicated when the needs of the customer cannot even be articulated objectively. For example, a study performed by Geymonat de Destefani and Whitefield discovered that one of the main methods used by individuals when choosing paint colors is "affective specification," in which individuals focus on emotional qualities (e.g. "comfortable" or "dramatic") and perceptual attributes (e.g. "light," "dark," "warm") rather than (relatively) more objective characteristics such as hue [4]. Definitions of these types of terms can vary considerably from person to person [5], so the task of determining optimal product colors from this kind of consumer feedback is reduced to educated guesswork at best.

A more objective means of working with consumer preferences can be found by using utility functions [6]. Utility measures the “attractiveness” of a given alternative to an individual in the form of a single scalar quantity. The amount of utility generated by a specific product can be represented as a function of the key attributes that define the product [7], making it possible to understand the relationship among attributes and identify worthwhile trade-offs [8]. Additionally, once utility functions have been determined for individual consumers, it is possible to apply clustering algorithms to the functional data to divide the population into market segments sharing similar preferences, allowing for optimal product designs to be developed for each market segment, thus increasing overall consumer satisfaction [9, 10].

One popular means of gathering preference information for a utility function is through the use of a conjoint analysis study. In conjoint studies, a product or line of products is first defined by a set of key measurable attributes. A parametric range is then determined for each of the key attributes, and a selection of levels from each range is chosen (usually both the high and the low end points, as well as one or more evenly spaced points in between). A set of products is created using every combination of these attribute levels and, if desired, this set of products can be reduced to a smaller fractional factorial subset of products (discussed in more detail in Section 3.4) before consumer preference data is collected.

One method of collecting this preference data is by asking each consumer to rank all of the potential product designs from most to least preferred, which is able to give information about both the desired levels for each of the attributes, as well as the importance of each attribute. For example, Page and Rosenbaum [9] discuss how the company Sunbeam used rankings based conjoint analysis to redesign their line of food processors. In this instance, the conjoint study first provided the company with information on the optimum levels for each of the 12 attributes in the study, but then showed that of these 12, only three were of major importance to the consumer, thereby giving designers and engineers significant freedom to alter the other attributes as needed in the redesign of the product.

Another popular method for collecting preference data is through ratings, in which each consumer will rate on a scale from 0 to 100, for example, his preference for each of the potential products. Moskowitz et al. [10] have successfully used conjoint ratings in determining consumers' preferences for pasta sauces. A benefit of conjoint ratings clearly illustrated by this study is that no data manipulation is necessary to allow comparison of preferences among the various levels of each attribute; the raw data in this case are preferences, greatly simplifying function generation on the researcher's end.

While both rankings and ratings based conjoint provide a wealth of information to the researcher, it has been shown that the quality of data received can be greatly reduced when the consumer feels mentally fatigued by the complexity of the tasks being presented [11]. In addition, these methods have been criticized for their lack of resemblance to consumers' actual behaviors while shopping [12].

More recently, however, Orsborn et al. [13] have developed a method for quantifying form preferences that uses choice based conjoint instead. Orsborn et al. were interested in measuring preference for vehicle forms, so the overall form of the vehicle was first "atomized," or broken into as many distinct, measurable component dimensions as possible. The seven most influential of these component dimensions were selected for manipulation, and an appropriate parametric range was identified for each. By combining maximum, minimum, and average values for each of these component dimensions, a large pool of potential vehicle designs was created. A selection of these vehicle designs were presented to consumers in sets of three, and the consumer was asked to choose his favorite from each set. Based on the frequency with which individuals chose the various designs, it was possible to construct a quadratic utility function for each of the seven attributes. By optimizing each equation, the most preferred component dimensions for each individual were discovered. These results were then verified through a follow-up choice survey with a 78% rate of success.

2.2. COLOR PREFERENCES

Research and experimentation in the area of color preferences has been going on since at least the 1890's [14], however conclusive results have not been reached over 100 years later. According to H.J. Eyesenck [14], early research focused primarily on three questions: (1) Is there a universal preference order for colors, or are preferences truly individual? (2) What are the differences in color preferences between men and women? (3) Is there a universal preference for saturated colors over unsaturated colors (i.e. tints and shades)? Answers to these each of these questions were often conflicting, and in 1940 H. J. Eyesenck attempted to resolve each of these issues with certainty (see [14] for a detailed list of early studies). To do this, he created a sample set of 10 colored paper swatches: six fully saturated hues (red, orange, yellow, green, blue, violet), three tints (red, orange, and green), and one shade (yellow). (The exact reasoning behind these specific choices is unknown.) Approximately 40 adults ranked these swatches from most to least favorite, and Eyesenck used these preference orders to, in his mind, definitively answer the three questions posed above. His research showed an objective preference order, with blue as the most preferred color. The only difference he found between preferences of men and women was in the lowest preference colors. Overall, men ranked the color yellow last and orange second to last; for women, these two rankings were reversed. Finally, he concluded that a secondary preference factor exists, separating those who prefer saturated colors from those who prefer tints and shades.

Unsatisfied with Eyesenck's conclusions, Guilford and Smith [15] created a much more complete work on the subject of color preferences in 1959. This work suggests that it is possible to predict the average preference level for a population based on the values of the color appearance attributes of hue, tint, and chroma, as determined by the Munsell color system [16].

In their work, a number of individuals viewed a total of 300 carefully chosen color specimens and rated the pleasantness of each on a 0 – 10 scale. Aggregating these responses, a set of isohedonic charts (what could be considered contour maps for preference) was created. Using these charts, one should be able to predict the average preference level of a population of men or women for any of ten different hues varying with regard to saturation and brightness.

More recently, Ou et al. [5] have had moderate success in predicting color preferences, using both color “emotions” and color appearance factors as predictor variables. In their study a total of 20 color swatches were chosen to represent a large range of hue, saturation, and lightness values. For each color swatch, an observer was presented with a pair of words representing opposite ends of an emotional spectrum (e.g. warm/cool, heavy/light, modern/classical, like/dislike) and was asked to select the word that more accurately described the color in question. The color samples were ranked on each of these emotional scales, based on the number of times they were described by the first word of each word pair, and each of these emotional frequency scales was compared to the overall like/dislike scale to determine if any statistical correlations existed [17]. In total, three different preference equations were developed, with (1) color emotions, (2) color-emotion factors, and (3) color appearance factors serving as the predictor variables. The predictive R^2 values for each of these equations ranged from 0.66 to 0.70, which is considerably better than random chance, an indication that the equations have successfully captured overall color preferences to some degree.

Overall, the prior works in preference modeling have proven the applicability of utility functions and choice surveys for mathematically modeling consumer preferences. Additionally, the existing research regarding color preferences provides support for the notion that preferences for color can be mathematically modeled as a function of measurable color attributes.

3. METHODOLOGY

This section will discuss the process used to create and verify utility functions for each individual, beginning with the selection of a product domain and proceeding through the selection of measurable attributes and a form for the utility function. Additionally, the methods used to collect preference data are discussed in detail.

3.1. SELECTING A PRODUCT DOMAIN

As previously mentioned, one of the limitations of nearly all existing color preference research is that subjects are asked to evaluate colors as stand-alone entities, separate from an product or application. This creates a rather significant logical problem, as one's preference for colors of automobiles, for example, is unlikely to be the same as his preference for kitchen appliances or sweaters [18]. Because color preferences are always specific to a given product, the first step of this research was to choose a type of product, or product domain, to which the sample set of colors would be applied.

In this case, backpacks were chosen to serve as the product domain for three reasons. First, backpacks can and do come in almost every conceivable color. This broad existing design space eliminates external constraints that would complicate the design of experiments.

Secondly, research has shown that color can play a more important role in purchase decisions when competing product choices are not considerably different from one another [18], as is the case with backpacks. In addition, consumers are less likely to choose from a limited set of "typical" colors for these types of lower risk purchases because advertisements are unlikely to have created any learned color associations. In short, a student's choice in backpack color is significant enough to involve some thought and emotion, but not so significant as to be practically predetermined by social norms.

Finally, backpacks are most regularly used by students, and since this research was conducted on a university campus, a ready supply of product consumers was available to serve as research test subjects.

3.2. CHOOSING A COLOR MODEL

For this research, it was necessary to first break the color one perceives into measurable components. This was done using the red, green, blue color model. The RGB color model is an additive model used to generate colors on electronic devices, such as televisions or computer screens. This model breaks perceived colors into red, green, and blue colored light components which can vary on an integer scale from 0 to 255. This model is called additive because darkness (that is, black) is produced when all three components are at their lowest level. In order to produce colors, light must be added, ultimately creating white when all components are at their highest levels (255). A shade of grey is produced when all three components are at the same level, and all remaining colors are produced by other combinations of level values.

3.3. SELECTING A FUNCTIONAL FORM

Utility functions can take any form, such as linear, quadratic, or exponential, though prior works by Chen et al. [19] and Moskowitz [20] have suggested that a quadratic utility function will accurately represent individual preferences for most applications, and the work of Orsborn et al. have successfully used quadratic utility functions to represent preferences for aesthetic form [13]. A quadratic function allows for a person's maximum preference for a given attribute to be at any point along the spectrum of possible values for that attribute. This is more flexible than a linear function which limits the maximum preference to either the high or low end of the spectrum.

However, Guilford's work with color preferences [15] (discussed in Section 2.2) produced multiple local maxima for preferences regarding hue, suggesting that quadratic equations might not be sufficient. For this reason, cubic utility functions were used instead, as shown in Equation 1, where x is the level of a given color component and a, b, c, and d are the coefficients for cubic regression.

$$u(x) = ax^3 + bx^2 + cx + d \tag{1}$$

It should be noted that this assumption will not have the effect of distorting preferences that are truly linear or quadratic, however, as those types of equations can simply be represented with zero coefficients for any unneeded higher order terms.

3.4. REDUCING THE DESIGN SPACE

Next, it was necessary to choose the specific colors for which preference data would be collected. In total there are 256^3=16,777,216 unique combinations of RGB values. Since this is clearly an unrealistically large number of sample products for an individual to evaluate, it was necessary to somehow reduce the design space to a more manageable without reducing the statistical reliability of the data that would be collected. To achieve this, a fractional factorial subset of the design space was used.

First, a smaller subset of values needed to be chosen out of the entire 0 – 255 parametric range for each of the three color attributes. In order to prevent the results from being biased in any way, it was necessary to choose evenly spaced values [12], the goal being to fairly represent the entire color space with as few samples as possible. This could be done with five evenly spaced levels per attribute, for a total of 5^3=125 colors in this reduced set. Thus, the levels used for the red, green, and blue color components were 0, 63, 127, 191, 255.

This set of 125 colors is referred to as the "full factorial" experiment design, because it includes all of the possible colors that can be made using the selected attribute levels. However, asking an individual to make even 125 observations would prove to be prohibitively time consuming, so fractional factorial methods were used to further reduce the design space.

In order to preserve the statistical validity of the experiment, the fractional factorial design would need to be both balanced and orthogonal [12]. In terms of experimental design, balance means that every level of every attribute appears in the experiment the same number of times. Using this research as an example, if level 0 of the Red attribute appears in the experimental set five times, levels 63, 127, 191, and 255 must also appear five times for the Red attribute to be balanced. Similarly, orthogonality means that every possible combination of two different attribute levels must appear in the experiment the same number of times.

The SAS software suite was utilized to determine the correct size of a fractional factorial that would meet these requirements, as well as the specific RGB levels for the colors contained in that set. The steps used are briefly summarized below, and the complete code is contained in Appendix A.

First, the %MKTRUNS macro was used to determine the possible sizes of balanced and orthogonal fractional factorial designs that could be made from the 125 color set. The macro showed that 100% D-efficient designs could be made from 25 or 50 elements. D-efficiency is a widely accepted mathematical measure of the goodness of an experimental design. In mathematical terms, it is a function of the geometric mean of the eigenvalues in the variance-covariance matrix [12], but as far as its application to this work, it is only important to understand that balanced and orthogonal design will always be 100% D-efficient.

The %MKTEX macro was then used to identify the 25 specific color samples that would be shown to consumers throughout the experiment. The colors, along with their RGB values, are shown in Table 3.1, below. (Note that RGB colors cannot be perfectly reproduced in print, so the color samples shown below do not appear exactly as they did in the study when they were viewed exclusively on computer screens.)

Table 3.1. Color Samples Used in Study

Red	Green	Blue	Swatch
0	0	0	
0	63	127	
0	127	255	
0	191	63	
0	255	191	
63	0	255	
63	63	63	
63	127	191	
63	191	0	
63	255	127	
127	0	191	
127	63	0	
127	127	127	

Red	Green	Blue	Swatch
127	191	255	
127	255	63	
191	0	127	
191	63	255	
191	127	63	
191	191	191	
191	255	0	
255	0	63	
255	63	191	
255	127	0	
255	191	127	
255	255	255	

3.5. COLLECTING PREFERENCE DATA

After the sample set of products has been determined, the next step is to begin collecting preference data from consumers. There are three main ways that this information can be gathered: ratings, rankings, or choices. As explained in Section 2.1, both ratings and rankings provide a great deal of information to the researcher, but place a large cognitive burden on the consumer, increasing the probability that the data will be distorted by fatigue [11]. A discrete choice survey, on the other hand, repeatedly presents the consumer with small subsets of the total set of products, and the consumer chooses his favorite from each of these subsets. In an efficient choice survey, each of the tasks is quite manageable, and results are less likely to be distorted.

In addition, neither rating nor ranking tasks accurately mimic the process a consumer typically uses when making purchase decisions. In reality, a consumer simply purchases the most preferred item and does not purchase the rest of the items [12], rarely giving dedicated thought to the preference order of the non-chosen alternatives. It is for these reasons that discrete choice surveys were used for data collection in this study.

In order to guarantee that the survey design remained balanced and orthogonal, the number of questions in the survey needed to be equal to the size of the fractional factorial used, or in this case, 25 questions. This survey length is reasonable, although so many of the same type of question in a row can become tedious, increasing fatigue risks [21]. In response, the survey was distributed online to be completed by respondents in their own time, allowing them to take breaks as needed. In addition, the questions in the survey were presented in a random order to each individual, spreading any fatigue or learning effects evenly throughout the survey.

3.6. CHOICE SURVEY DEVELOPMENT

To design the choice survey, SAS's %CHOICEFF macro was used. The survey consisted of 25 questions, each with three choices. Because the survey was both balanced and orthogonal, as discussed previously, this meant that the each of the 25 colors contained in the fractional factorial set appeared three times throughout the survey. The final design was evaluated with the %MKTEVAL macro to confirm that it was both balanced and orthogonal.

The survey design from SAS was then translated into pictorial survey questions, like the one shown below in Figure 3.1 and was distributed online through commercially available software provided by SurveyGizmo. The complete survey design is provided in Appendix B.

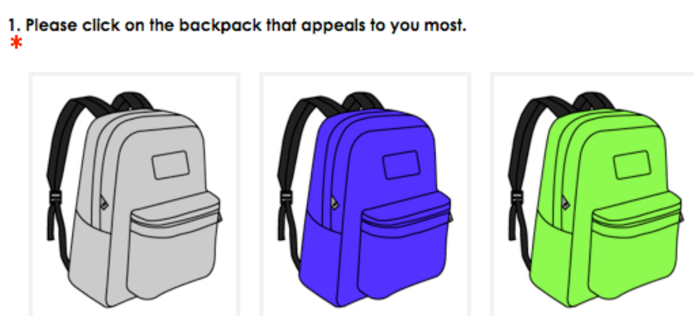

Figure 3.1. Example Survey Question

3.7. GENERATING UTILITY FUNCTIONS

In most marketing applications, these coefficients are found by using software to apply a logit or probit model to aggregate consumer data [12]. However, Hazelrigg [22] proved that while these aggregate methods can produce acceptable equations for predicting market demand, they are likely to generate erroneous preference models. For this reason, it is necessary to evaluate utility functions on a consumer-specific basis. However, stable partworth coefficients cannot be found in a choice based survey logit or probit methods without aggregating results from many respondents [23].

In order to determine partworth utilities for individuals, then, it was necessary to use Luce's Choice Axiom [24]. This method is based on probability of choice, as shown by Equation 2, which states that the probability of an object being chosen is equal to the weight of that object (w_i) divided by the sum of the weights of all the objects from which the choice was made (w_j). In this case, the probability that a given color component level is chosen by an individual is equal to the number of times a design including that level is chosen in the survey, divided by the number of times it appeared.

$$P(i) = \frac{w_i}{\sum_j w_j} \tag{2}$$

For example, the question shown previously in Figure 3.1 is shown again in Figure 3.2, this time with the color component levels for each of the choices. If a respondent chose the blue backpack in the middle, the totals for Red 64, Green 0, and Blue 255 would each increment by one.

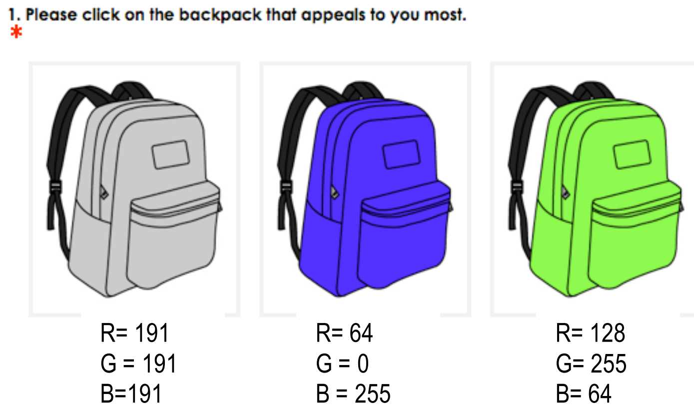

Figure 3.2. Example Survey Question with Color Component Levels

At the end of the process, a table similar to the one shown in Table 3.2 would be created. Dividing each of these totals by the number of times each level was seen, in this case 15, the partworth utilities are found, also seen in Table 3.2. Each value represents the probability that the consumer will choose a design containing the corresponding level for that attribute. As a result, the partworth utilities can range from 0 to 1, with higher values indicating a higher preference.

Table 3.2. Example Choice Totals and Partworth Utilities

	Choice Totals			Partworth Utilities		
Level	**Red**	**Green**	**Blue**	**Red**	**Green**	**Blue**
0	6	6	6	0.400	0.400	0.400
64	6	3	8	0.400	0.200	0.533
128	5	5	4	0.333	0.333	0.267
191	3	7	4	0.200	0.467	0.267
255	5	4	3	0.333	0.267	0.200

Cubic utility functions were then fit to the scatter plots of the partworth values, as shown in Figure 3.3. The maximum value of each of these functions occurs at the color component level that the individual prefers most. The respondent shown in Figure 3.3 prefers a low value of red (37), a higher level of green (195), and a low value for blue (46).

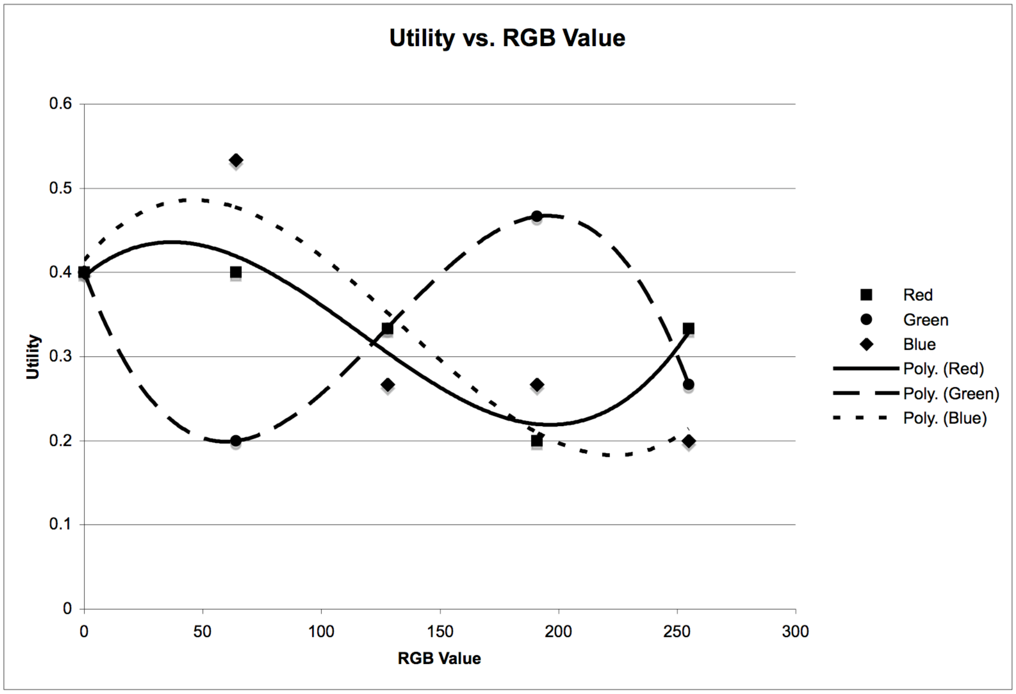

Figure 3.3. Utility Functions for Example Respondent

Additionally, Luce's Choice Axiom assumes that a consumer's overall utility is represented as a summation of his utility for the each of the individual attributes, as indicated in Equation 3, where $u(x_i)$ is the utility of an individual color component.

$$U = \sum_{i=1}^{n} u(x_i) \tag{3}$$

The advantage of this assumption is that it allows each attribute utility function to be optimized individually, meaning standard derivative based optimization can be used, without the need for more complex computer algorithms. However, this assumption is limited in that it forces the preferences for each individual attribute to be unrelated to preferences for any other attributes. This assumption, and its implications for color, will be discussed in detail in Section 4.

3.8. CREATING PRODUCT DESIGNS

Finally, these equations can be used to create high utility colors for each individual. Under these assumptions, the highest utility color is made up of each of the most preferred color component levels. For the respondent shown above, the highest utility color is shown in Figure 3.4, below.

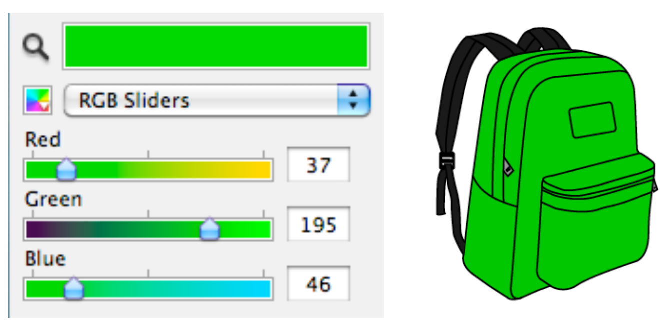

Figure 3.4. Highest Utility Backpack Color for Example Respondent

4. CONSUMER STUDIES

This section utilizes the methodology from Section 3 in two real world consumer studies. By studying the results of the first of these studies, several potential improvements to the methodology were identified. In the second study, these changes were implemented alongside the original methodology, allowing for a comparison of results.

4.1. INITIAL STUDY

4.1.1. Data Collection. Using the methods described in Sections 3.1 – 3.6, a 25 question choice survey was developed regarding preferences for backpack colors. The survey was distributed online to 78 students in a freshmen-level engineering class. The pool of respondents consisted of 64 men and 14 women, ranging in age from 18 to 32, with more than 50% of the respondents being either 18 or 19 years old.

After the surveys had been completed, a utility function was developed for each individual, as explained in Section 3.7. To test the validity of these utility functions, a two question follow-up survey was created for each individual. In the first question, the individual was given a choice among three different backpack colors, all generated using his specific utility function. Of the three colors, one was the individual's highest utility color, one was his lowest utility color, and one had an average, or neutral, utility value. If the utility function accurately captured color preferences, it would be expected that individuals would choose the highest utility color an overwhelming majority of the time and the lowest utility color very infrequently.

The second question in the follow-up survey asked whether the choice that was made in Question 1 was the individual's (a) favorite backpack color, (b) favorite backpack color out of the three colors that were shown, or (c) a random choice because none of the color choices were appealing. This question was included to better understand the degree of success of the utility functions. For example, if many of the respondents chose randomly, it would be erroneous to conclude the model was an accurate predictor of preference, even if all of them chose the highest utility color.

4.1.2. Results. Of the 78 students who completed the initial survey, 55 also completed the follow-up. The breakdown of their responses can be seen in Figure 4.1, below.

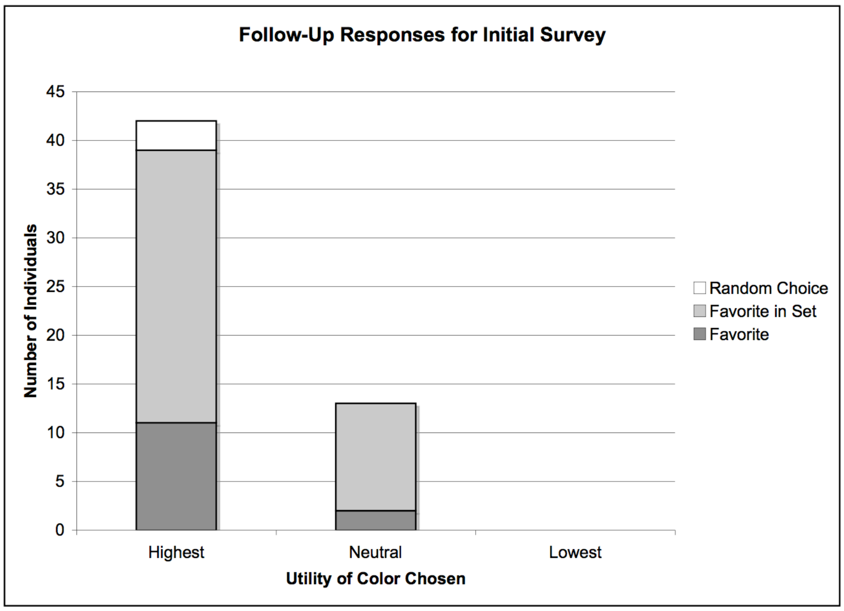

Figure 4.1. Summary of Follow-Up Responses to Initial Survey

The majority of individuals (76.4%) chose the highest utility color. The remaining 23.6% of individuals chose the neutral utility color, and no one chose the lowest utility color. However, the majority of respondents also indicated that the chosen color was only their favorite of the three colors that were shown, and not their overall favorite color. This information suggests that, though the utility functions are correlating higher utilities with higher preferences quite successfully, the methodology is not completely capturing preferences with the present form of the utility function, so it was necessary to explore other potential forms of the utility function.

4.2. ATTRIBUTE UTILITY FUNCTIONS

The overall form of a utility function consists of two parts: 1) the functional form of the individual attribute utilities and 2) the method with which those attribute utility functions are combined to form the overall utility. The data gathered in the initial survey was used to look deeper into both of these facets of the overall utility function.

As explained previously, the utility functions for the red, green, and blue color attributes were developed by fitting a cubic curve to the partworth utilities that were determined by Luce's Choice Axiom. While acceptable results were achieved with this functional form, the possibility remained that a simpler functional form could fit the data just as well, or that another form could fit the data even better.

In order to explore these possibilities, various potential functional forms were fit to the data and the resulting R^2 values were compared. Statistically, R^2 is a measure of the amount of variance that is removed when a given function is used to relate the predictor variables to the output variable. R^2 values can range from 0 to 1, with 1 meaning the function fit the data perfectly and 0 indicating a complete lack of fit. In other words, a higher R^2 value indicates a stronger relationship between the individual's stated preferences and the function being used to approximate those preferences.

Another factor to consider is the fact that increasing the order of a polynomial regression will never cause R^2 to decrease. If the addition of extra terms provides only a minimal reduction in variance, however, they are typically not included in the final form of the equation. For example, if a cubic utility function fits the data only marginally better than a quadratic, it would be advisable to switch to the less complex functional form.

Table 4.1 shows the R^2 values achieved using cubic, quadratic, and linear functional forms. As can be seen, the cubic functions did fit the data substantially better than either of the other more simple functional forms. For that reason, the cubic functional form was maintained throughout the rest of the studies.

Table 4.1. Variance Reduction Using Various Attribute Level Utility Functions

Form	R^2
Cubic	0.812
Quadratic	0.696
Linear	0.398

4.3. COMBINATION OF ATTRIBUTE UTILITIES

One key area for improvement in the current model is to take interaction effects into account when determining overall color preferences. Interaction effects refer to the effect that combinations of two or more different color component levels appearing together has on the overall utility function. This is most clearly seen in the following example. Shades of grey are represented in the RGB color scale by all three color components having equal levels (e.g. R=100, G=100, and B=100). If an individual has a strong preference for the color grey, but not necessarily for a certain shade of grey, their preference would not be for any specific color component levels, as long as the levels are all equal to one another. This interaction has no way of being captured in the current model because each color components are considered independently without consideration for the possible effects of such interactions.

Ideally, the utility function would be able to fully incorporate all possible interaction effects. However, predicting coefficients for interaction variables would first require that a larger fractional factorial, and therefore larger survey, be used [12], further increasing the risk of erroneous results due to fatigue. Additionally, a method other than Luce's Choice Axiom would need to be used to develop the utility functions. As explained in Section 3.5, this would require that preference data be gathered from ratings or rankings based conjoint, rather than a discrete choice survey, or that utility functions be developed for an aggregate sample of consumers instead of for individuals. Since both of these options were undesirable for reasons previously discussed, two hypothetical solutions were proposed.

4.3.1. Multiplicative Utility. The first potential solution to this problem was to consider the overall utility function to be a product of the attribute level utilities, rather than a sum, as shown in Equation 4, below. When this multiplication is carried out for the three attribute utility functions, interaction terms, and the corresponding coefficients, are created. A partial example of this new equation is shown for the domain of RGB color in Equation 5, below, where $a_1 - a_n$ are the coefficients, and *R, G,* and *B* are the numeric values of the red, green, and blue attributes levels.

$$U = \prod_{i=1}^{n} u(x_i) \tag{4}$$

$$U = a_1R + a_2G + a_3B + a_4RG + \cdots + a_nR^3G^3B^3 \tag{5}$$

4.3.2. Utility of Greys. For the second potential solution, the attempt was made to capture preferences for one of the more significant interactions, the one that results in shades of grey, in its own utility function. This interaction was chosen based on discussions with respondents which indicated that these preferences make up a significant portion of backpack color preferences. Essentially, an additional "grey" utility function would be developed using the same method as the utility functions for the red, green, and blue color attributes. The partworth **utilities** for each of the five shades of grey contained in the survey would be found using Luce's method, simply the number of times the grey was chosen, divided by the number of times it appeared in the survey. To complete the utility function, a cubic curve would then be fit to the data points.

This new function would not be summed or multiplied with the other three to determine the overall utility of a color, however. Instead, overall utility would now be a piecewise defined function, as seen in the additive form in Equation 6, and in the multiplicative form in Equation 7.

It should be noted that in the additive form of the equation, the grey utility is multiplied by 3 to account for the fact that it is taking the place of all three color components and must therefore be weighted accordingly. For the same reason, the grey utility is cubed in the multiplicative form of the equation.

$$U = \begin{cases} 3 * u(grey) & if\ red = green = blue \\ u(red) + u(green) + u(blue) & all\ else \end{cases} \tag{6}$$

$$U = \begin{cases} u(grey)^3 & if\ red = green = blue \\ u(red) * u(green) * u(blue) & all\ else \end{cases} \tag{7}$$

To develop the grey utility function, choices for greys could be accounted for in one of two ways. In Grey Method 1, these choices would be counted in addition to the regular choices for the red, green, and blue attributes. In other words, the red, green, and blue utility functions would be unchanged, and a separate function would be used to calculate the utility in the specific instance when all three color attributes existed at the same level.

In Grey Method 2, however, each choice in the survey is considered as a choice for a color *or* a choice for grey. This change in the way choices are counted affects the choice totals for each individual, and therefore results in different utility functions for all four of the attributes. For the sake of comparison, Table 4.2 contains the partworth utilities for one individual using the original method, as well as both forms of grey handling.

Table 4.2. Comparison of Partworth Utilities Using Different Grey Handling Methods

	Choice Totals				Partworth Utilities						
					No Grey Handling			Grey Method 1			
Level	Red	Green	Blue	Grey	Red	Green	Blue	Red	Green	Blue	Grey
0	6	6	6	3	0.400	0.400	0.400	0.400	0.400	0.400	1.000
64	6	3	8	2	0.400	0.200	0.533	0.400	0.200	0.533	0.667
128	5	5	4	2	0.333	0.333	0.267	0.333	0.333	0.267	0.667
191	3	7	4	2	0.200	0.467	0.267	0.200	0.467	0.267	0.667
255	5	4	3	2	0.333	0.267	0.200	0.333	0.267	0.200	0.667

	Choice Totals				Partworth Utilities			
					Grey Method 2			
Level	Red	Green	Blue	Grey	Red	Green	Blue	Grey
0	3	3	3	3	0.250	0.250	0.250	1.000
64	4	1	6	2	0.333	0.083	0.500	0.667
128	3	3	2	2	0.250	0.250	0.167	0.667
191	1	5	2	2	0.083	0.417	0.167	0.667
255	3	2	1	2	0.250	0.167	0.083	0.667

4.4. FURTHER DATA ANALYSIS

The two new potentially beneficial changes to the functional form can be combined into a total of six different new utility functions, as follows:

1. Addition
2. Multiplication
3. Addition (Grey Method 1)
4. Multiplication (Grey Method 1)
5. Addition (Grey Method 2)
6. Multiplication (Grey Method 2)

While these various methods make sense in theory, a preliminary test of validity was performed on the existing data from the initial study. To do this, the utility of each of the choices in the original survey was calculated for each individual using each of the six different utility functions, above. Next, the highest utility color for each question was compared with the color the individual actually chose. A utility function that is accurately representing an individual's preference should be able to successfully recreate the choices in the original survey. In an ideal situation, a second set of data would have been used for verification, however this method was deemed to serve as an acceptable preliminary filter.

The overall results of this test, by percentage of choices modeled correctly, can be seen below in Table 4.3. The simplest form of the utility function, addition with no grey handling, was used as a benchmark, and only functional forms that performed as well or better, as determined through a difference of means hypothesis test, were considered for further evaluation. This filtering method removed addition, using grey handling method 1, but passed through the remaining 5 forms.

Table 4.3. Preliminary Rates of Success for Different Utility Function Forms

	No Grey Handling	Grey Method 1	Grey Method 2
Addition	74.60	69.35	75.25
Multiplication	74.10	71.95	76.50

4.5. EXPANDED STUDY

4.5.1. Survey Distribution. Finally, in order to test the validity of these five potentially useful functional forms, a new study was performed. The first part of the survey, in which consumer preferences were gathered, was unchanged from the previous version. A total of 291 students in a freshman-level engineering class participated in this research, and the demographic breakdown of this sample was similar to that in the initial study. This sample of respondents consisted of 215 men and 76 women, ranging in age from 18 to 40. Again, more than 90% of the respondents were 21 or younger, and half were ages 18 and 19.

4.5.2. Follow-Up Questions. The key difference in this study, however, is the follow-up survey. In the initial study, one form of the utility function (a simple linear function of the cubic attribute utilities) was used to generate one multiple choice question for each individual, in which the highest, lowest, and neutral utility backpack colors were compared. The expanded study, on the other hand, tested all five potential forms of the utility function (as identified in the previous section), using not one, but five questions for each, for a total of 25 questions. Using multiple questions for each method would decrease the impact of "false positive" responses (that is, responses in which the highest utility color was chosen randomly by an indifferent respondent) and serve to better reveal the true success of each of the methods.

The method for developing the colors for these follow-up questions was more complicated for two reasons. First, to find the maximum or minimum utility color meant optimizing each of the attribute utility functions individually, then combining each of those optimal attribute values into one color. In other words, there is only one maximum and one minimum utility color. However, as soon as the search extends beyond either pole of the utility spectrum, there can be several different sets of RGB values that generate the same utility for the individual. Standard stochastic search optimization methods can get stuck in local optima, reducing the likelihood that they will find a diverse variety of solutions. Additionally, all of the utility functions that incorporated some sort of grey preference handling were piecewise defined and discontinuous, meaning that no standard search-based optimization methods could be used.

For these reasons, exhaustive enumeration was used to generate the colors for the follow-up surveys. The outputs are then sorted by value, making it easy to find multiple solutions within any given range of utilities. This method is beneficial for the fact that it works exactly the same way for any type of utility function, no matter how complex. The drawback, however, is that exhaustive enumeration is very computationally intensive and can become quite time consuming, even with modern computing capabilities. Due to the discontinuous nature of the utility functions being considered, this was the only option that could guarantee that no optimal (or non-optimal) solutions were missed.

In practice, high utility colors were those that ranked within the top 10% of all colors generated. In other words, these colors had a utility of 90 – 100% of the maximum possible utility. Neutral utility colors were those with 45 – 55% of the maximum utility, and low utility colors had only 0 – 10% of the maximum utility. The range of 10% was chosen because it was a large enough window that it actually generated five non –identical colors within each range, but small enough that each of the three colors in a given question had distinctly different utilities.

As illustrated in Figure 4.2, five colors were chosen at even intervals from within each of these categories. Additionally, the colors shown in each question were pulled from the same "slot" within each category, as indicated by the circled elements in Figure 4.2.

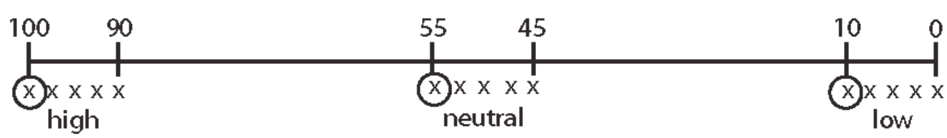

Figure 4.2. Utility Relationship of Colors Used in Follow-Up Questions

4.5.3. Results of Follow-Up Surveys. The individual-specific follow-up surveys were also distributed online and were completed by 256 of the original 291 respondents. The results of the survey are summarized in Table 4.4.

Table 4.4. Percentage of Choices, by Utility Method

Functional Form	High	Neutral	Low
Addition	74.31	17.63	8.06
Multiplication	74.15	18.58	7.27
Addition (Grey 2)	70.36	18.42	11.23
Multiplication (Grey 1)	67.91	16.13	15.97
Multiplication (Grey 2)	67.67	16.84	15.49

As the table shows, both of the methods that did not incorporate grey handling performed essentially the same as one another, and superior to the methods that did attempt to take grey preferences into account. Figures 4.3 – 4.7 contain this data graphically, including standard deviations. As seen in the graphs, below, only the values for the high utility choice are statistically significantly, as the standard deviation bars overlap for the neutral and low choice percentages in all of the methods.

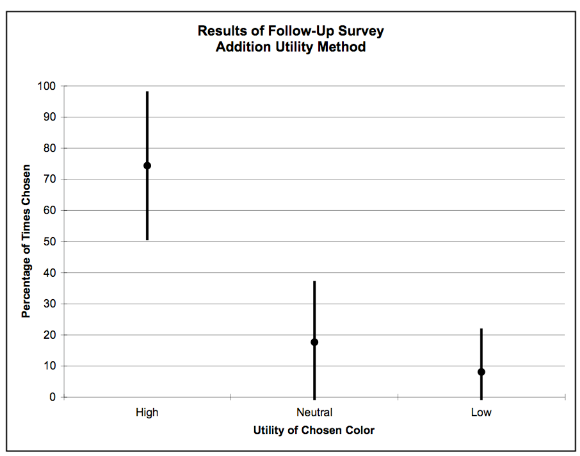

Figure 4.3. Follow-Up Results – Addition

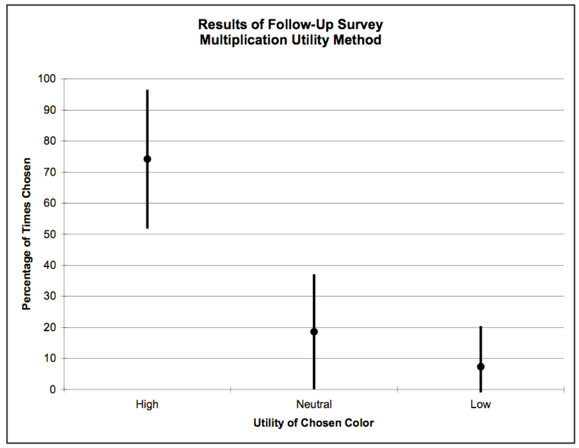

Figure 4.4. Follow-Up Results – Multiplication

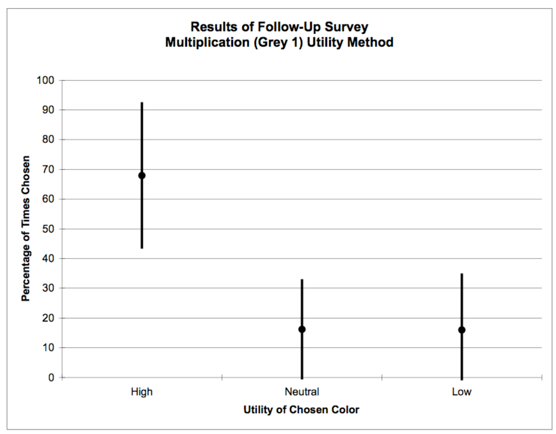

Figure 4.5. Follow-Up Results – Multiplication (Grey Method 1)

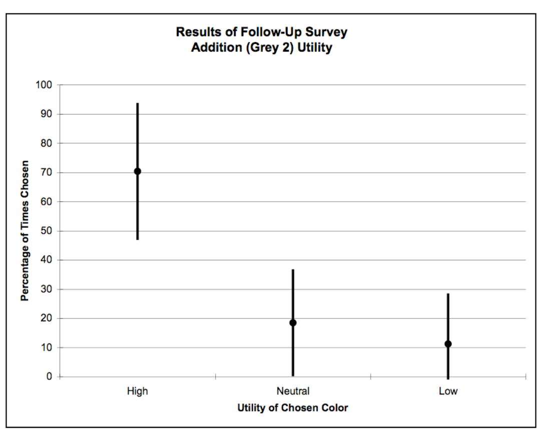

Figure 4.6. Follow-Up Results – Addition (Grey Method 2)

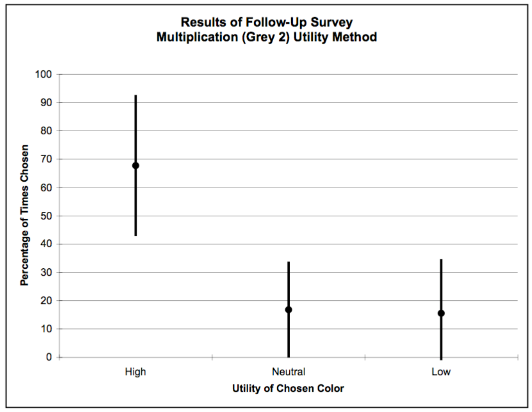

Figure 4.7. Follow-Up Results – Multiplication (Grey Method 2)

While none of the utility functions that incorporated grey handling methods performed better than the benchmark method, there was substantial variance in the performance of these methods from question to question. In particular, each of these methods performed exceptionally well in the questions where the highest utility color was one of the three options (i.e. the question that would be produced with the circled elements in Figure 4.2). A comparison of these results can be seen in Figure 4.8 below.

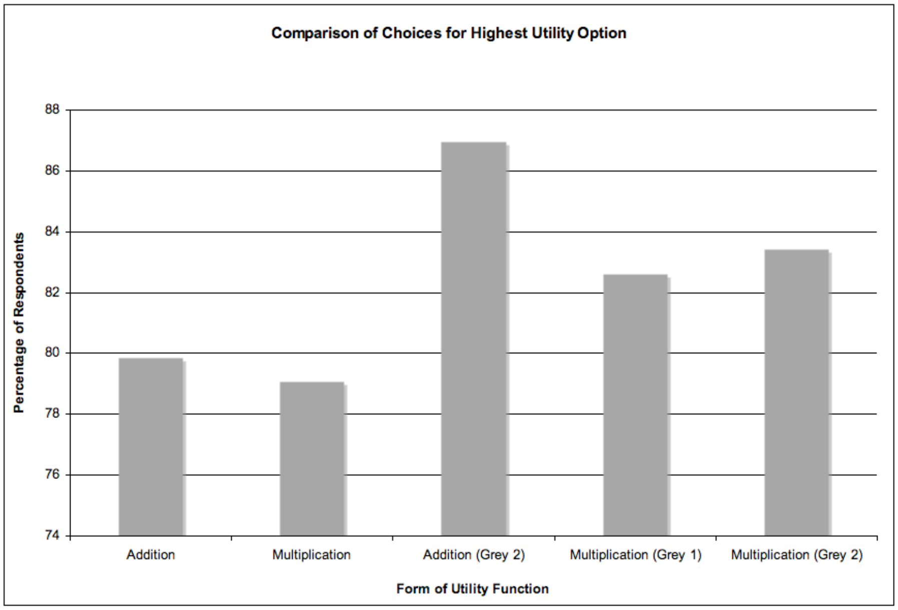

Figure 4.8. Comparison of Results for Questions Containing the Highest Utility Color

In particular, the addition functional form, using grey handling method 2 had the largest percentage of individuals choosing the highest utility option, as well as the smallest percentage of individuals choosing the low utility option. Why these methods failed to maintain this high rate of success in questions offering lower high utility options is unknown, but their success in this specific instance is not without merit. These results indicate that when it comes to determining *the* optimum color, a method that accounts for grey preferences is essential. On the other hand, when a range of very good options is required instead, a method without grey handling is preferable.

5. CONCLUSIONS AND FUTURE WORK

While this study has developed a foundational methodology for representing colors with utility functions, this body of research can be enriched with additional work in any of several areas.

A valuable first step in future work would be to test this method using a different product space. It is possible that while backpack color preferences can be somewhat successfully represented by their red, green, and blue color components, preferences for other products might not follow suit. Additionally, the methodology has only been verified using a sample drawn from a relatively young, predominantly male population of engineering college students. It would be interesting to see if similar results are obtained using a more broadly representative sample of individuals.

Next, it will be necessary to accurately and fully account for all interaction effects in one continuous function. The partial success of the grey handling methods employed in this research indicates that interactions are important. However, these methods failed to produce superior results in any situation where the maximum utility option was not one of the available choices. A more reliable functional form must be developed, which will likely require that individual preferences be gathered through either a ratings or rankings based conjoint method, which have been successfully used for these purposes elsewhere [9, 10, 20]. In addition, the final utility form should be continuous, such that optimization methods can be applied for more efficient evaluation. Exhaustive enumeration is an acceptable academic approach, but it is too time consuming and computationally intensive to be used in real-world applications.

Finally, after these problems have been solved to some degree of completion, it would be pertinent to address product designs that involve multiple colors. In the fairly simple product space of backpacks, individuals repeatedly commented that their actual favorite product color would be “red with black accents” or “black with blue and green stripes.” Prior research [17] suggests that preferences for color pairings cannot easily be associated to preferences for individual colors, however it makes sense to resolve the simpler task before moving forward. Though the problem is different, the methods developed in this research on single color preferences could certainly be adapted to multi-color situations, and perhaps even combined with previously discussed form-preference work [13] to incorporate pattern preferences as well.

In conclusion, the translation of aesthetic preferences to objective functions is a complex task, and this research has responded by outlining a methodology and providing substantial preliminary verification to guide future researchers as they seek to refine and build upon the existing body of research.

APPENDIX A

COMPLETE SAS CODE FOR SURVEY GENERATION

```
/*Generates a fractional factorial experiment with 25 designs for

/*an experiment with three attributes, each at five levels.


     %mktex(5 5 5, n=25);
     proc print; run;


/*Evaluates fractional factorial design to confirm balance and
/*orthogonality.

     %mkteval;


/*Creates a 25 question choice survey.

     %mktlab(data=design, int=f1-f3);
     proc print; run;

     %choiceff(data=final, bestout=sasuser.survey25,
     model=class(x1-x3), nsets=25, maxiter=10, flags=f1-f3,
     options=nodups, beta=zero);

     proc print; by set; id set; run;


/*Evaluates survey design to confirm balance and orthogonality.
     %mkteval;
     proc print;  run;
```

APPENDIX B

COMPLETE SURVEY DESIGN

Question	Attribute Levels		
	Red	Green	Blue
1	191	191	191
	64	0	255
	128	255	64
2	0	0	0
	191	64	255
	64	128	191
3	0	64	128
	255	255	255
	191	128	64
4	64	0	255
	0	64	128
	191	191	191
5	255	128	0
	191	64	255
	0	255	191
6	191	255	0
	64	64	64
	128	191	255
7	64	255	128
	191	128	64
	128	64	0
8	255	255	255
	191	0	128
	128	64	0
9	128	255	64
	64	191	0
	255	64	191
10	191	0	128
	255	64	191
	128	191	255
11	191	128	64
	0	64	128
	128	0	191
12	255	255	255
	64	128	191
	128	64	0
13	0	255	191
	255	0	64
	128	191	255
14	64	64	64
	255	128	0
	128	0	191
15	255	191	128
	191	255	0
	128	0	191
16	64	255	128
	191	191	191
	0	0	0

Question	Attribute Levels		
	Red	**Green**	**Blue**
17	128	128	128
	64	191	0
	0	255	191
18	64	128	191
	191	64	255
	255	191	128
19	0	191	64
	191	255	0
	128	128	128
20	255	0	64
	64	255	128
	0	128	255
21	128	255	64
	0	128	255
	191	0	128
22	0	191	64
	255	128	0
	64	0	255
23	64	191	0
	0	128	255
	255	0	64
24	0	0	0
	255	191	128
	64	64	64
25	0	191	64
	255	64	191
	128	128	128

BIBLIOGRAPHY

[1] Hauser, John R., and Clausing, Don, 1988, "The House of Quality," Harvard Business Review, pp. 63-73.

[2] Liu, Yili, 2003, "Engineering Aesthetics and Aesthetic Ergonomics: Theoretical Foundations and a Dual-Process Research Methodology," Ergonomics, 46(13/14), pp. 1273-1292.

[3] Bacon, Frank R., and Butler, Thomas W., 1981, Planned Innovation, University of Michigan Press, Ann Arbor.

[4] Geymonat de Destefani, Lucila R., and Whitfield, T. W. Allan, 2008, "Esthetic Decision-Making: How Do People Select Colours for Real Settings?," Color Research and Application, 33(1), pp. 55-60.

[5] Ou, Li-Chen, Luo, M. Ronnier, Woodcock, Andree, and Wright, Angela, 2004, "A Study of Colour Emotion and Colour Preference. Part I: Colour Emotions for Single Colours," Color Research and Application, 29(3), pp. 232-240.

[6] von Neumann, John, and Morgenstern, Oskar, 1944, Theory of Games and Economic Behavior, Princeton University Press, Princeton.

[7] Ben-Akiva, Moshe, and Lerman, Steven R., 1985, Discrete Choice Analysis: Theory and Application to Travel Demand, The MIT Press.

[8] Thurston, Deborah L., 1991, "A Formal Method for Subjective Design Evaluation with Multiple Attributes," Research in Engineering Design, 3(2), pp. 105-122.

[9] Page, Albert L., and Rosenbaum, Harold F., 1987, "Redesigning Product Lines with Conjoint Analysis: How Sunbeam Does It," Journal of Product Innovation Management, 4, pp. 120-137.

[10] Moskowitz, Howard R., Jacobs, Barry E., and Lazar, Neil, 1985, "Product Response Segmentation and the Analysis of Individual Differences in Liking," Journal of Food Quality, 8(2/3), pp. 169-181.

[11] Swait, Joffre, and Adamowicz, Wiktor, 2001, "The Influesnce of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching," Journal of Consumer Research, 28, pp. 135-148.

[12] Kuhfeld, Warren F., 2005, Marketing Research Methods in SAS, SAS Institute Inc., Cary.

[13] Orsborn, Seth, Cagan, Jonathan, and Boatwright, Peter, 2009, "Quantifying Aesthetic Form Preference in a Utility Function," Journal of Mechanical Design, 131(6).

[14] Eysenck, H. J., 1941, "A Critical and Experimental Study of Colour Preferences," American Journal of Psychology, 54, pp. 385-394.

[15] Guilford, J. P., and Smith, Patricia C., 1959, "A System of Color-Preferences," American Journal of Psychology, 72, pp. 487-502.

[16] 1929, Munsell Book of Color, Munsell Color Company, Baltimore.

[17] Ou, Li-Chen, Luo, M. Ronnier, Woodcock, Andree, and Wright, Angela, 2004, "A Study of Colour Emotion and Colour Preference. Part Iii: Colour Preference Modeling," Color Research and Application, 29(5), pp. 381-389.

[18] Grossman, Randi Priluck, and Wisenblit, Joseph Z., 1999, "What We Know About Consumers' Color Choices," Journal of Marketing Practice, 5(3), pp. 78-88.

[19] Chen, Wei, Wiecek, Margaret M., and Zhang, Jinhuan, 1999, "Quality Utility - a Compromise Programming Approach to Robust Design," Journal of Mechanical Design, 121(2), pp. 179-187.

[20] Moskowitz, Howard, 1998, "On Fitting Equations to Sensory Data: A Point of View, and a Paradox in Modeling and Optimizing," Journal of Sensory Studies, 15, pp. 1-33.

[21] Bozorgzadeh, Amir, 2008, "Respondent Fatigue: Online Survey Tactics," Vue, pp. 24-26.

[22] Hazelrigg, G. A., 1996, "The Implications of Arrow's Impossibility Theorem on Approaches to Optimal Engineering Design," Journal of Mechanical Design, 118, pp. 161-164.

[23] Huber, Joel, Wittink, Dick R., Johnson, Richard M., and Miller, Richard, 1992, "Learning Effects in Preference Tasks: Choice-Based Vs. Standard Conjoint," Proc. Sawtooth Software Conference.

[24] Luce, R. Duncan, 1959, Individual Choice Behavior: A Theoretical Analysis, John Wiley and Sons, New York.

VITA

Hannah L. Turner received her B.S. in Interdisciplinary Engineering from the Missouri University of Science and Technology in Rolla, Missouri in August 2008. In May, 2010 she received her M.S. degree in Manufacturing Engineering from the Missouri University of Science and Technology. She has published a conference paper pertaining to this research. In addition, Hannah L. Turner was inducted into Kappa Mu Epsilon Honor Society in 2005.