Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2013
August 4-7, 2013, Portland, Oregon, USA

DETC2013-12622

TOWARDS UNDERSTANDING THE ROLE OF INTERACTION EFFECTS IN VISUAL CONJOINT ANALYSIS

Brian Sylcott
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA, 15213
Email: sylcott@cmu.edu

Jeremy J. Michalek
Mechanical Engineering
Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA, 15213
Email: jmichalek@cmu.edu

Jonathan Cagan
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA, 15213
Email: cagan@cmu.edu

ABSTRACT

This work investigates how attribute interactions affect consumer preference for product form attributes in choice based conjoint analysis. Interaction effects are present when preference for the level of one attribute is dependent on the level of another attribute. When some or all interaction effects are negligible, a factional factorial experimental design can be used to reduce the number of variants shown to the respondent and keep the survey size manageable. This is particularly important when the presence of many parameters or levels makes full factorial designs intractable. However, if there is uncertainty about the independence of attributes or if interaction effects are relevant, a fractional factorial design may create biased estimates and misleading conclusions. This work examines the role of interaction effects in visual conjoint analysis, an extension to traditional conjoint analysis that allows for product form attributes that vary continuously. Many visual conjoint analysis studies assume interaction effects are negligible. We conduct preliminary tests on this assumption in three visual conjoint analysis studies. The results suggest that interactions can be significant or negligible in visual conjoint, depending on both consumer preferences and shape parameterization. We suggest that randomized designs are generally better than fractional factorial designs at avoiding errors due to the presence of interactions and/or the organization of profiles into choice sets.

INTRODUCTION

The ability to capture and characterize consumer preference is an important tool for design engineers. When developing a design solution, engineers need useful information about their target users so that they can better tailor their designs. Preference modeling is a common way to gather this information. One of the most popular methods for constructing consumer preference models is conjoint analysis [1]. In conjoint analysis direct feedback is solicited from consumers in the form of product surveys. These surveys present participants with multiple product profiles chosen to span the design space without conflating the effects of attributes [2]. In the past decade, conjoint has been applied in a variety of engineering design contexts [3–5]. Although originally used to characterize consumer preference for the functional attributes of a product, over time the method has evolved and extensions have been added to accommodate aesthetic attributes. Orsborn et al. presented one of those extensions, visual conjoint analysis [6]. In that work visual conjoint analysis was used to show aesthetic preference for complex designs such as the front of vehicles could be captured through conjoint analysis. The designs were the composition of several Bezier curves whose control points were varied to create alternative designs. The parameterization in visual conjoint analysis studies allows for a product's shape to vary continuously through the design space.

One of the main considerations when planning out a conjoint study is the number and makeup of the survey questions. In order to develop an effective survey researchers employ design of experiments [7]. Design of experiments refers to a set of rules that determine how to efficiently plan out experiments in order to answer specific questions. The most comprehensive design (in a discrete space) is the full factorial [8]. In a full factorial design all the possible combinations of attribute levels are represented. For example a full factorial study design containing 5 attributes each with 3 levels requires 3^5 or 243 product profiles. Participant responses to a full factorial survey design allow for the calculation of both main effects-the effect a product attribute level has on preference for the entire product-and interaction effects-the effect one attribute level has on preference for another. This information can be used to generate a utility function that describes participant product preference. Unfortunately, as the number of questions required for a full factorial design scales exponentially with the number of attributes, the load on the respondents quickly becomes too great. As a result, it is often impractical to use full factorial surveys to model preference for complicated products that have many attributes. Fractional factorial designs enable researchers to estimate specific effects with fewer questions when other effects are believed to be negligible. This results in a smaller survey size; however, if the assumed-negligible effects are in fact significant, estimates will be biased [9].

It is important to note that factorial design theory is based on linear designs and is not directly applicable to choice based tasks. In conjoint analysis the linear design is analogous to a ratings task, but a choice task is distinct because the dependent variable (utility) is not observed directly. Traditional design of a choice based task typically begins with developing a balanced and orthogonal linear design and organizing the profiles arbitrarily into choice sets [10]. Alternative methods for optimizing the survey design exist when priors are available on the model parameters [11–13], or when adaptive methods are available [14] but we do not pursue those here.

In this work, we use case studies to investigate the validity of ignoring interaction effects for conjoint analysis cases that involve aesthetic attributes. Our goal is to get a better understanding of how the inclusion or exclusion of interactions can impact the performance of consumer preference models of product form constructed through visual conjoint analysis.

PREVIOUS WORK

There are many examples of researchers using conjoint analysis and other methods to model consumer preference for form attributes. Swamy et al. successfully used conjoint analysis to model preference for vehicle headlight shape [15]. The product representation in that work was the outline of the headlight composed of Bezier curves. Headlight shape was also the subject of preference modeling in separate work completed by Petiot and Dagher [16]. Here an alternative method to conjoint analysis is shown to be able to capture preference. This method uses multidimensional scaling to build a perceptual space that yields interpretable perceptual dimensions. Reid et al. used a visual conjoint method to quantify the relationship between aesthetics and perceived environmental friendliness [17]. That work had participants rate two-dimensional vehicle silhouettes on environmental friendliness. The results showed that cars with smoother curves were more likely to be thought of as being inspired by nature while boxier cars were less likely. Macdonald et al. [18] used conjoint analysis to model semantic messages of wine flavor associated with different wine bottle shapes. The flavor models were used in conjunction the shape models to optimize wine portfolio characteristics such as flavor, quantity produced, and profitability. Kelly and Papalambros [19] presented a method for optimizing a product based on aesthetic preference data and engineering performance characteristics. Conjoint analysis is used to capture preference for the shape of a beverage bottle. Shape preference was then plotted along with shape dependent engineering characteristics to create a Pareto front that illustrated the tradeoffs between aesthetic form preference and actual functional performance.

In work by Tseng et al. a methodology was presented for capturing preference for stylistic attributes [20]. The subject of that work was a vehicle design represented by line drawing silhouettes. Here neural networks were used to capture preference and genetic algorithms were used to create optimal designs based on preference. Sylcott et al. [21] used the same vehicle representation scheme as Tseng et al. in a visual conjoint analysis study. The work involved developing a meta-conjoint approach in order to combine form and function preference. In work by Turner et al. a conjoint framework was used to model preference for color [22]. The attributes in that study were the red, green, and blue components that make up the color of a backpack and the levels were the component intensity values. That work showed that preference for something as subjective color can be captured in a utility function.

In many of these cases and others that deal with modeling preference for subjective characteristics, interaction effects were assumed to be negligible and were not included in the preference models. Missing interactions can bias parameter estimates and result in misleading conclusions. However, there is evidence that in some situations the added complexity that comes with including interaction effects is not extremely beneficial and that the variance in the term estimation can negatively impact the model performance [23].

Green and Srinivasan suggested that interaction effects are important in situations that involve surveying sensory phenomena, styling, and aesthetic features [24]. The literature on multi-attribute utility theory has also emphasized the importance of nonlinear effects [25–28]. In the following case studies this work investigates examples of preference models with and without interaction effects in order to get a better understanding of how interaction effects influence the performance of visual conjoint based models.

APPROACH

This work consists of three different case studies. In each study utility functions are developed based on the results from choice based visual conjoint surveys. When making a choice between alternatives, Eq. 1 defines the total utility associated with alternative j out of J total alternatives:

$$\begin{aligned} u_j &= x_j\beta + \varepsilon_j \\ &= v_j + \varepsilon_j \end{aligned} \quad . \tag{1}$$

Here, u_j is the total utility associated with the j^{th} design alternative, x_j contains the attribute values for the design alternative (and their combinations, such as interactions), and β is a vector of unknown regression parameters. The quantity $x_j\beta$, or v_j accounts for the observable portion of the alternative's utility while ε_j accounts for the unobservable portion, which is treated as a random variable. In this work maximum likelihood estimation (MLE) is used to solve for β.

Multinomial logit (MNL) models relate the utility of a design and its alternatives to the probability of the focal design being chosen [29]. β is found by maximizing the probability that the model will generate the observed data. Assuming each of N total decision makers makes their selections independently and that the error term from Eq. 1 follows an extreme value distribution, this likelihood, L can be expressed as Eq. 2 [30]:

$$L = \prod_j P_j^{n_j}, \tag{2}$$

where n_j is the number of respondents that choose alternative j and the logit probability, P_j is defined by Eq. 3 [31]:

$$P_j = \frac{e^{v_j}}{\sum_1^M e^{v_{km}}}. \tag{3}$$

Eq. 3 describes the probability of choosing alternative j from the set of M alternatives present in a choice set k. The log likelihood, LL, is obtained by taking the log of Eq. 2 resulting in Eq. 4:

$$LL = \sum_j n_j \ln\left(P_j\right). \tag{4}$$

An optimal value for β can be found by maximizing Eq. 4 (a monotonic transformation of Eq. 2) with respect to β. Once β is specified the utility model is complete and can then be evaluated for performance.

The metrics chosen to evaluate the performance of the utility models in this work are the log likelihood, LL, equivalent average likelihood (EAL), hit rate (HR), and mean absolute share error (MASE), all calculated at the optimal β.

EAL (equivalent average likelihood) simply normalizes LL with respect to the size of the data set by assessing the common likelihood necessary at each data point to produce an equivalent overall likelihood for the model. EAL is calculated using Eq. 5:

$$EAL = L^{\frac{1}{N}}. \tag{5}$$

HR (hit rate) is a measure of how well a utility function can predict consumer choice. It is calculated by comparing the observed selections with the predicted selections for each choice set [32]. The predicted selections are determined by calculating the utility associated with each choice option using the attribute β coefficients. The choice option with the highest utility is the predicted choice for that set. Each time the observed selection matches the predicted choice counts as a hit-otherwise it is a miss. The hit rate is calculated with Eq. 6:

$$HR = \frac{1}{N \cdot K} \sum_1^K n_h, \tag{6}$$

where K is the total number of choice sets and n_h is the number of respondents who selected the highest predicted P_j for each choice set k.

MASE (mean absolute share error) is used to evaluate how well the observed market shares line up with the predicted market shares. It is calculated by taking the average of the absolute difference between the observed and predicted market share for each design alternative as shown in Eq. 7 [33]:

$$MASE = \frac{\sum_{j=1}^{J} \left| s_{j,PRED} - s_{j,OBS} \right|}{J}, \tag{7}$$

where $s_{j,PRED}$ is the predicted market share, and $s_{j,OBS}$ is the observed market share. Together these metrics are used to evaluate the overall performance of the utility models in each of the following case studies.

CASE STUDIES

Case 1: Line Angle Preference Illustration

Methodology

This example is used to illustrate the role model formulation plays in determining the importance of interactions and their impact on model performance. In order to make this illustration as clear as possible a simple model was developed using simulated response data. Taking this approach we created responses consistent with a particular preference. This allowed for a cleaner evaluation of the model of that preference.

Visual conjoint studies often use outlines and silhouettes to represent product shape. That representation is simplified further in this example as preference for the angle of a line is modeled. The subject of the example is a straight line drawn between two points, p_1 and p_2, whose coordinate positions vary. The x-coordinates remain constant and each y-coordinate has three possible values: p_y = {-1,0,1}. The survey consists of 2 attributes each with 3 levels leading to a total of 3^2 or 9 possible profiles. In each question there are three alternative profiles. An example trial is shown in Figure 1.

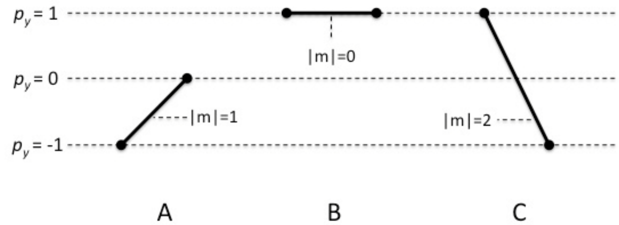

Figure 1: Example line Angle Trial

All possible combinations of the design profiles are used to develop 84 choice sets (choice sets involving identical profiles were ignored). For each trial, synthetic data from 102 subjects were generated to simulate respondents who have a preference for flat lines and select alternatives with the smallest slope. The number 102 was chosen because it is divisible by 2 and by 3 allowing the number of respondents to be split evenly when options appear with the same slope. The data were generated using the following rules: (1) if the magnitude of the slope of one of the alternatives is smaller than the other two, the alternative is chosen by 100 simulated respondents and each of the other two is chosen by 1 respondent; (2) if the magnitude of all the slopes of all three alternatives is the same, each alternative is chosen by 34 of the simulated respondents; (3) if there is a tie for the lowest slope magnitude, each of the tied alternatives is selected by 50 of the simulated respondents and the third alternative is selected by 2 respondents. Since ln(0) is undefined each option always had at least 1 respondent.

By adapting Eq. 1 the observed utility, v, for all design alternatives can be represented by Eq. 8:

$$v = \mathrm{X}\beta, \tag{8}$$

where **X** is the j x i coded design matrix, and i is the number of regression parameters. Each row of **X** corresponds to a design alternative. For a given design matrix, the utility associated with the j^{th} alternative is described by Eq. 9:

$$v_j = \sum_{1}^{i} \beta_i x_{ij} \tag{9}$$

The composition of **X** is determined based on which β coefficients are to be estimated. Referring back to Figure 1, the design matrix representing this trial would be $\mathrm{X} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$.

The first row of **X** corresponds to alternative A. p_{1y} is at level 1 and p_{2y} is at level 2. Rows 2 and 3 of **X** correspond to alternatives B and C, respectively. For part-worth models dummy variables are used to code the design matrix so that each column corresponds to a part-worth level of an attribute. Only n-1 levels are required to estimate part-worth values for each attribute. Level 2 was chosen to be the base level and omitted. Using this coding transforms **X** into a 3 x 4 matrix, $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$. Multiplying out Eq. 9 leads to Eq. 10:

$$v_j = \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j}. \tag{10}$$

Here x_{11} and x_{13} correspond to the two levels of attribute 1 (other than the base level 2) while x_{21} and x_{23} correspond to the two levels of attribute 2. When attribute i is at level 1 x_{i1} = 1 otherwise x_{i1} = 0. Accounting for interactions in this model requires the addition of four columns to the design matrix, the product of the elements of columns 1 and 3, 1 and 4, 2 and 3, and 2 and 4. Now, **X** is a 3 x 8 matrix. The additional columns in **X** transform Eq. 10 into Eq. 11:

$$\begin{aligned} v_j &= \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j} \\ &+ \beta_5 x_{11j} x_{21j} + \beta_6 x_{11j} x_{23j} \\ &+ \beta_7 x_{13j} x_{21j} + \beta_8 x_{13j} x_{23j} \end{aligned} \tag{11}$$

In the initial model formulation the attributes were taken to be the y-coordinates of the endpoints, p_1 and p_2. However, there is more than one way to define a line. Since we have prior knowledge that preference is based on the magnitude of the slopes of these lines, it is reasonable to define the lines using their starting point, p_1 and slope magnitude $|(p_{2y}-p_{1y})/1|$.

Each of the two attributes in this formulation will also have 3 levels, p_y = {-1,0,1} and m = {0,1,2}, so the structure of Equations 10 and 11 will remain the same; however, **X** will not. Going back to the trial in Figure 1, the initial **X** will have a different second column, $X = \begin{bmatrix} -1 & 1 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}$. This difference carries through the coded matrices.

Results - All choice sets

For each of the two formulations two MNL models were created. These models are summarized in Table 1.

Table 1: Summary of choice models

Model	Formulation	Main Effects	Interaction Effects
Part-worth	End Point	A (Eq. 10)	B (Eq. 11)
	Point Slope	C (Eq. 10)	D (Eq. 11)

The solution to each of these models was found using the modified quasi-Newton method implemented in Matlab's *fminunc* function. The part-worth model results presented in Table 2 are based on simulated responses to all 84 possible choice sets. Since each possible combination is present in the model, there are no holdout questions.

Table 2: Line angle part-worth results using all choice sets

β		A	B	C	D
x_{11}	0	-0.22 (0.03) ***	-3.98 (0.11) ***	-0.00 (0.04)	-0.00 (0.07)
x_{13}	0	-0.22 (0.03) ***	-3.98 (0.11) ***	-0.00 (0.04)	-0.00 (0.07)
x_{21}	0	-0.22 (0.03) ***	-3.98 (0.11) ***	3.98 (0.09) ***	3.98 (0.10) ***
x_{23}	0	-0.22 (0.03) ***	-3.98 (0.11) ***	-2.94 (0.12) ***	-1.96 (253.24)
$x_{11}*x_{21}$	0		7.95 (0.19) ***		0.00 (0.09)
$x_{11}*x_{23}$	0		1.04 (0.20) ***		-0.98 (253.24)
$x_{13}*x_{21}$	0		1.04 (0.20) ***		0.00 (0.09)
$x_{13}*x_{23}$	0		7.95 (0.19) ***		-0.98 (253.24)
In Sample					
LL	-9412.91	-9339.93	-3600.65	-3600.65	-3600.65
EAL	33%	34%	66%	66%	66%
HR	39%	39%	77%	77%	77%
MASE	33%	32%	1%	1%	1%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$

The first column in Table 2 lists the attribute effects. For each model the estimated coefficients are listed with the standard error in parentheses and asterisks to indicate statistical significance. The table also shows in-sample performance metrics. Higher LL and EAL values indicate a higher probability that the model is a good fit for the data. HR will only reach 100% if all of the respondents choose the highest utility option in all the choice sets. That will not happen in these examples because many choice sets include multiple options with the same utility and respondents split their decisions over those options. As MASE approaches 0% so does the difference between the observed and predicted market shares. We include for reference a null model, in which the utility of all alternatives is taken as zero (no information). The null model is a baseline that indicates how the models would perform if they were fit at random.

With the endpoint formulation, the solution to Model A, which lacks interaction effects is $\beta x_{11} = \beta x_{13} = \beta x_{21} = \beta x_{13} = -0.22$. This solution indicates that on average there is a disadvantage in moving from level 2 to levels 1 or 3 for both attributes. The only information gained is that the best alternative is for both attributes to be at level 2. The model performs no better than the null model. Because preference is based on each endpoint's position relative to the other, main effects alone are not sufficient to model preference. In the solution to Model B the main effects are $\beta x_{11} = \beta x_{13} = \beta x_{21} = \beta x_{13} = -3.98$. However, the interaction terms in this model give additional descriptive capability. For example, $\beta x_{11}*x_{21}$ is considerably greater than $\beta x_{11}*x_{23}$. Plugging these values into Eq. 11 shows that if p_{1y} is at level 1 it is much better for p_{2y} to be at level 2, (v = -3.98) than at level 3 (v = -6.91). The best option would be for p_{2y} to be at level 1 as well (v = 0). This matches well with a preference for the line to be as flat as possible. As expected, the inclusion of interaction terms allows Model B to perform much better on all of the performance metrics than model A. This was not case for the point-slope formulation.

Unlike the endpoint formulation, the point-slope formulation includes the slope as a main attribute. This attribute describes the relationship between the positions of p_{1y} and p_{2y} and allows the formulation to model preference without including interaction terms. As a result, in this formulation Model C supplies all the information necessary to model preference. $\beta x_{11} = \beta x_{13} = 0$ and is not statistically significant, indicating that there is no preference for the starting point of the line. The values of βx_{21} and βx_{23} indicate the best slope is at level 1 (value of zero) no matter the position of the starting point. This matches preference for a line with a slope of 0. The performance of Model C in the point-slope formulation is identical to that of Model B. The performance of Model D is identical to Model C. None of the interaction terms are statistically significant showing there is nothing gained by including them.

The key observation here is that when the shape is modeled with one representation (endpoint), interactions are critical, whereas when the same shape is modeled with a different representation (point-slope), interactions are negligible. Thus, the question of what role interactions play in visual conjoint depends not only on the shape and preferences for that shape, but also on the way the shape is parameterized. This parameterization is a choice that researchers make when using these types of representations in visual conjoint studies.

Results - 9 choice sets

It is often impractical and unnecessary to gather subject responses to all possible choice set combinations. Instead, a subset of the choice sets are chosen and organized into an efficient survey. The information obtained from different surveys is not always equivalent. One common metric for evaluating the goodness of a survey is D-efficiency. The formula often used to calculate D-efficiency is shown in Eq. 12 [10]:

$$D = 100 \frac{1}{N \left| \left(X' X \right)^{-1} \right|^{\frac{1}{P}}}. \qquad (12)$$

Here, **X** is an orthogonally coded design matrix of size N x p. A survey that is orthogonal will have a D-efficiency equal to 100%. However, this formula is based on a linear survey design and only takes into account which design profiles are included in the survey, not how they are organized into choice sets. This example was also used to explore how the choice of survey design can affect results.

In a linear design the minimum number of runs needed to develop an orthogonal design to estimate main and interaction effects for this example is 9. The SAS software package is used to generate a 100% efficient survey with 9 choice sets. MLE is used to estimate β for the simulated responses to this survey. 9 questions were also generated as a holdout sample. The in sample and hold out part-worth results are listed in Table 3.

Table 3: Line angle part-worth results 9 choice sets

	β	Null Model	A	B	C	D
	x_{11}	0	-0.15 (0.08) *	-4.08 (0.47) ***	0.05 (0.14)	-0.00 (0.19)
	x_{13}	0	-0.15 (0.08) *	-4.08 (0.47) ***	0.05 (0.14)	-0.00 (0.19)
	x_{21}	0	-0.15 (0.08) *	-4.08 (0.47) ***	4.21 (0.30) ***	4.08 (0.36) ***
	x_{23}	0	-0.15 (0.08) *	-4.08 (0.47) ***	-2.59 (0.37) ***	-1.72 (*)
	$x_{11}*x_{21}$	0		8.31 (0.65) ***		0.16 (0.31)
	$x_{11}*x_{23}$	0		1.50 (1.06)		-0.86 (*)
	$x_{13}*x_{21}$	0		1.50 (1.06)		0.16 (0.31)
	$x_{13}*x_{23}$	0		8.31 (0.65) ***		-0.86 (*)
In Sample	LL	-1008.53	-1003.58	-344.32	-344.52	-344.32
	EAL	33%	34%	69%	69%	69%
	HR	35%	37%	80%	80%	80%
	MASE	24%	33%	1%	1%	1%
Hold Out	LL	-1008.53	-993.75	-579.65	-579.49	-579.65
	EAL	33%	34%	53%	53%	53%
	HR	42%	40%	62%	62%	62%
	MASE	36%	24%	1%	1%	1%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$, (*) SE with complex roots

Overall the models built from the 9-question survey performed as well as the models built from all possible choice sets. The values of the regression coefficients are similar and each model's performance metrics were comparable to those found using all choice sets. As before, the metrics for models B, C, and D are far better than the null case both in and out of sample. Additionally, the same relationship between the models and their formulations is present.

There are hundreds of thousands of orthogonal surveys that can be developed with 9 choice sets. While traditional choice designs based on linear models organize a factorial of profiles into choice sets arbitrarily [34], when the modeler has priors for coefficients, optimal design methods exist to determine the best choice sets for estimating the parameters with the lowest error [13], but we do not pursue optimal design approaches here. 1,000 surveys were generated by randomly assigning a full factorial design into choice sets three separate times. This prohibits any design profile from appearing in a choice set more than once. Responses to each survey were simulated and models were constructed. The same 9 holdout questions were used to evaluate each model. Table 4 shows the mean of the performance metrics with standard deviations.

Table 4: Line angle mean performance 9 choice sets (n=1000)

		A	B	C	D
In Sample	LL	-955 (50)	-386 (121)	-389 (121)	-386 (121)
	EAL	35% (2%)	66% (9%)	66% (9%)	66% (9%)
	HR	46% (9%)	76% (9%)	76% (9%)	76% (9%)
	MASE	30% (6%)	1% (0%)	1% (0%)	1% (0%)
Hold Out	LL	-1079 (136)	-633 (66)	-596 (41)	-626 (50)
	EAL	31% (3%)	50% (3%)	52% (2%)	51% (3%)
	HR	37% (7%)	60% (5%)	61% (4%)	60% (5%)
	MASE	26% (2%)	5% (3%)	2% (3%)	5% (3%)

The mean performance metrics for models B, C, and D are far better than the null model. Other than the log likelihood the performance metrics have relatively small standard deviations. This suggests that many of the models have similar performance. However, when comparing individual surveys, performance can vary greatly. The model A results from one survey showed HR to be 62% and MASE to be 1% while another resulted in an HR = 51% and MASE = 24%. The D-efficiency of both of these designs was equal to 100%. These results suggest that the manner in which design profiles are organized into choice sets can have a significant impact on model performance and that D-efficiency alone may not be sufficient to find the best survey design. Conducting randomized surveys, where each respondent receives a different group of choice sets, can mitigate error associated with arbitrary assignment of profiles into conjoint questions.

The main objective of this example is to illustrate how, in visual conjoint, interactions are sometimes critical and other times negligible. Identifying whether they matter in a given case requires some exploration and as shown will depend on the formulation of the model. In the endpoint formulation, interaction effects were required to perform better than the null model, which is equivalent to random guessing. The inclusion of interaction terms leads to a substantial improvement in performance over all the metrics. In the point-slope formulation, models performed just as well with or without interactions. The results from this study serve as motivation for researchers to further investigate the role of interaction effects as they can have significant impact on model performance. These results also show how models based on the same data can perform better or worse based on how the model is formulated.

Case 2: Vase Shape Preference Illustration
Methodology

The subject of this example is the outline of a flower vase. The vase consists of the four Bezier curves depicted in Figure 2. The control points of the Bezier curves are parameterized so that three attributes of the vase vary. These attributes are the height to average width ratio, the top to bottom width ratio, and the curvature of the sides. Symmetry is enforced between curves 2 and 4. Each attribute has three levels. The attributes are depicted in Figure 3.

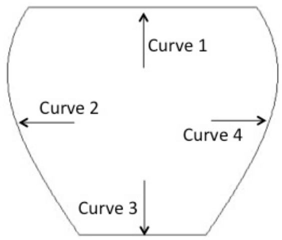

Figure 2: Four Bezier curves that make up vase model

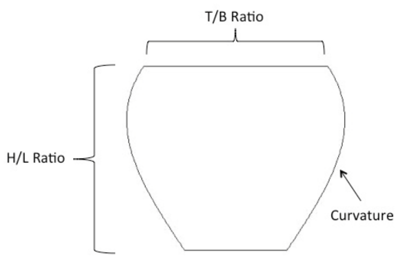

Figure 3: Vase attribute

The three attributes, each with three levels, yield 27 possible profiles. As done in the previous example all combinations of these 27 profiles were combined in to choice sets. Taking three alternatives per set yields 2925 possibilities (choice sets involving identical profiles were ignored). An example trial is depicted in Figure 4.

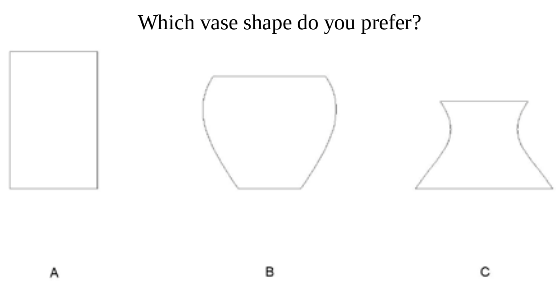

Figure 4: Example Vase preference trial.

For each trial, synthetic data from 102 subjects was generated to simulate respondents who have a preference for a square vase. When each of the three attributes is at level 2 the vase is a square. Preference for the design alternative decreased as the model deviated from the ideal. This deviation is equal to $\sum_{1}^{3} |x_i - 2|$. The data were generated using the following rules: (1) if the deviation of one of the alternatives is smaller than the other two, the alternative is chosen by 100 simulated respondents and each of the other two is chosen by 1 respondent; (2) if the deviation of all of the three alternatives is the same, each alternative is chosen by 34 of the simulated respondents; (3) if there is a tie for the lowest deviation, each of the tied alternatives is selected by 50 of the simulated respondents and the third alternative is selected by 2 respondents. This formulation presumes that interactions are not important to the respondent.

As before, the utility for each alternative *j* is determined by Eq. 9. In this case only part-worth models are used. The three-three level attributes result in a 3 x 6 dummy coded design matrix for the main effects model corresponding to Eq. 13:

$$v_j = \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j} + \beta_5 x_{31j} + \beta_6 x_{33j} . \qquad (13)$$

Here, x = {height to average width ratio, top to bottom width ratio, the side curvature}. Accounting for first order interactions leads to a 3 x 18 design matrix and Eq. 14:

$$v_j = \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j} + \ldots + \beta_{17} x_{23j} x_{31j} + \beta_{18} x_{23j} x_{33j} . \qquad (14)$$

Results - All choice sets

As before, the solution to each of these models was found using the modified quasi-Newton method implemented in Matlab's *fminunc* function. The part-worth model results are presented in Table 5.

Table 5: Vase part-worth results all choice sets

	β	Null Model	Eq. 13	Eq. 14
	x_{11}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	x_{13}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	x_{21}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	x_{23}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	x_{31}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	x_{33}	0	-3.58 (0.01) ***	-3.73 (0.03) ***
	$x_{11}*x_{21}$	0		0.11 (0.03) ***
	$x_{11}*x_{23}$	0		0.11 (0.03) ***
	$x_{11}*x_{31}$	0		0.11 (0.03) ***
	$x_{11}*x_{33}$	0		0.11 (0.03) ***
	$x_{13}*x_{21}$	0		0.11 (0.03) ***
	$x_{13}*x_{23}$	0		0.11 (0.03) ***
	$x_{13}*x_{31}$	0		0.11 (0.03) ***
	$x_{13}*x_{33}$	0		0.11 (0.03) ***
	$x_{21}*x_{31}$	0		0.11 (0.03) ***
	$x_{21}*x_{33}$	0		0.11 (0.03) ***
	$x_{23}*x_{31}$	0		0.11 (0.03) ***
	$x_{23}*x_{33}$	0		0.11 (0.03) ***
In Sample	LL	-327770.98	-124348.48	-124334.76
	EAL	33%	66%	66%
	HR	29%	78%	78%
	MASE	32%	1%	1%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$

The results from the part-worth model indicate that the best design alternative has all attributes at level 2. There is a decrease in utility associated with moving from level 2 in either direction for all attributes. Although interactions were found to be significant, the EAL, HR, and MASE for the main effects model are identical to the first order interaction model suggesting there is no need to include interaction effects.

Results - 27 choice sets

For this example, the minimum number of runs needed to develop an efficient linear survey capable of estimating main and interaction effects is 27. SAS was used to generate a 100% D-efficient 27-question survey. MLE is used to estimate β for the simulated responses to this survey. 27 additional questions were also generated as a holdout sample. The in sample and holdout part-worth results are listed in Table 6.

Table 6: Vase part-worth results 27 choice sets

	β	Null Model	Eq. 13	Eq. 14
	x_{11}	0	-3.20 (0.14) ***	-6.26 (0.75) ***
	x_{13}	0	-3.25 (0.14) ***	-6.59 (0.75) ***
	x_{21}	0	-3.23 (0.14) ***	-6.59 (0.71) ***
	x_{23}	0	-3.22 (0.14) ***	-6.28 (0.79) ***
	x_{31}	0	-3.26 (0.13) ***	-7.06 (0.68) ***
	x_{33}	0	-3.24 (0.14) ***	-6.78 (0.67) ***
	$x_{11}*x_{21}$	0		2.68 (0.69) ***
	$x_{11}*x_{23}$	0		2.41 (0.72) ***
	$x_{11}*x_{31}$	0		2.81 (0.71) ***
	$x_{11}*x_{33}$	0		2.63 (0.81) ***
	$x_{13}*x_{21}$	0		2.43 (0.85) ***
	$x_{13}*x_{23}$	0		2.44 (0.98) **
	$x_{13}*x_{31}$	0		3.47 (0.64) ***
	$x_{13}*x_{33}$	0		3.21 (0.64) ***
	$x_{21}*x_{31}$	0		3.25 (0.61) ***
	$x_{21}*x_{33}$	0		3.20 (0.66) ***
	$x_{23}*x_{31}$	0		3.12 (0.67) ***
	$x_{23}*x_{33}$	0		2.90 (0.59) ***
In Sample	LL	-3025.58	-1175.12	-1150.14
	EAL	33%	65%	66%
	HR	47%	79%	79%
	MASE	30%	1%	1%
Hold Out	LL	-3025.58	-1290.81	-1293.61
	EAL	33%	63%	63%
	HR	26%	76%	76%
	MASE	29%	1%	2%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$

The results from the SAS designed survey were comparable to the results from the survey of all choice sets. Both models performed well above the null model and there was not a major difference in performance between the main and interaction effects models.

An additional 1,000 27-question surveys were generated by randomly assigning a full factorial design into choice sets three separate times. Design profiles were prohibited from appearing in a choice set more than once. Responses to each survey were simulated and models were constructed. The same 27 holdout questions were used to evaluate each model's out of sample performance. Table 7 shows the mean of the performance metrics with standard deviations.

Table 7: Vase mean performance 27 choice sets (n = 1000)

		Null Model	Eq. 13	Eq. 14
In Sample	LL	-3025.58	-1144 (184)	-1138 (185)
	EAL	33%	66% (4%)	66% (4%)
	HR	47%	78% (5%)	78% (5%)
	MASE	30%	1% (0%)	1% (0%)
Hold Out	LL	-3025.58	-1303 (7)	-1326 (32)
	EAL	33%	62% (0%)	62% (1%)
	HR	26%	76% (0%)	76% (0%)
	MASE	29%	1% (0%)	3% (1%)

The results in Table 7 indicate there was a relatively small variation in performance between the difference surveys and on average each survey performed reasonably better than the null model.

Results - 9 choice sets

Although some interaction terms were found to be significant in the previous sections the models in this example performed comparably with and without interaction effects both in and out of sample. In this situation ignoring interactions will not lead to decreased predictive capability. However, doing so allows for the use of a main effects design. The minimum number of runs needed to develop an efficient survey to estimate the main effects in this example is 9. SAS is again used to develop an efficient survey. The results from this survey are listed in Table 8.

Table 8: Part-worth results from 9 question survey

	β	Null Model	Eq. 13
	x_{11}	0	-3.46 (0.26) ***
	x_{13}	0	-3.19 (0.29) ***
	x_{21}	0	-3.27 (0.24) ***
	x_{23}	0	-3.27 (0.30) ***
	x_{31}	0	-2.98 (0.29) ***
	x_{33}	0	-3.29 (0.25) ***
In Sample	LL	-1008.53	-226.91
	EAL	33%	78%
	HR	37%	91%
	MASE	38%	2%
Hold Out	LL	-3025.58	-1313.00
	EAL	33%	62%
	HR	26%	76%
	MASE	29%	3%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$

The regression coefficients and the holdout sample performance metrics listed in Table 8 are consistent with those found in the previous sections.

Another 1,000 9-question surveys were generated by randomly assigning a single full factorial design into choice sets. Responses to each survey were simulated and models were constructed. The same 27 holdout questions were used to evaluate each model's out of sample performance. Table 9 shows the mean of the performance metrics with standard deviations.

Table 9: Vase mean performance 9 choice sets (n = 1000)

	β	Null Model	Eq. 13
In Sample	LL	-1008.53	-379 (103)
	EAL	33%	0.67 (7%)
	HR	37%	0.78 (8%)
	MASE	38%	0.01 (0%)
Hold Out	LL	-3025.58	-1364 (133)
	EAL	33%	0.61 (3%)
	HR	26%	0.76 (1%)
	MASE	29%	0.04 (2%)

The results listed in Table 9 show that on average, over the sample set the main effects models performed much better than the null model. Additionally, because the interaction effects were negligible in the simulated preference data, a fractional factorial main effects design was sufficient to estimate the model and provide good predictions.

Case Study 3: Vehicle Shape Preference
Methodology

The first two cases were simple examples designed to illustrate how interaction effects can be either critical or negligible and how modeling decisions can impact their importance. The first two cases were extremes on two opposite ends of the spectrum. A real world case would most likely fall somewhere in between. It is often difficult to know the extent that interaction effects matter in a given situation before hand. Single surveys that are capable of estimating all main and interaction effects are often too large to be practical. An alternative approach is to use a random survey design for each participant. With a large enough sample size, random surveys are capable of estimating both main and interaction effect. This approach was followed in this case.

In case 3 surveys are issued to real participants in order to capture their preference for vehicle shape. The subject of this study is a vehicle design depicted by line drawing silhouettes built using the scheme developed by Tseng et al. [20]. An example of the vehicle representation is shown in Figure 5.

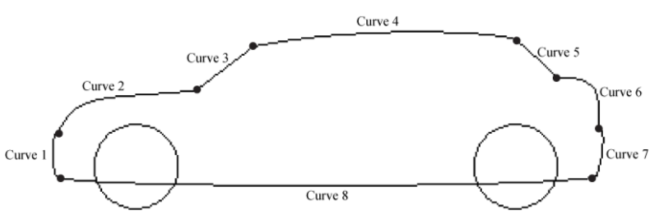

Figure 5: Eaxmple vehicle

As shown in Figure 5, these representations are the composition of eight Bezier curves. The control points of the curves are parameterized in a method that allows several major features of the design to be varied continuously. In this study four attributes were varied, the ground clearance, body height, hood length, and trunk length. These attributes are depicted in Figure 6.

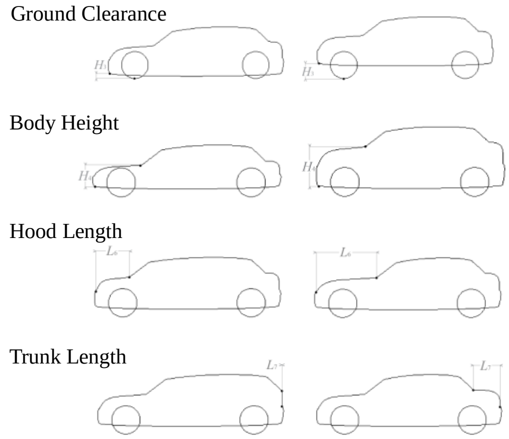

Figure 6: Vehicle attributes

Each of the four attributes consisted of three levels. Levels 1 and 3 of each attribute are depicted in Figure 6. Level 2 is midway between the values shown in Figure 6. There were a total of 81 possible designs. The Sawtooth software package was used to organize the profiles in to random designs for each participant. Each survey question presented three design alternatives for the participant to choose from. A sample trial is shown in Figure 7.

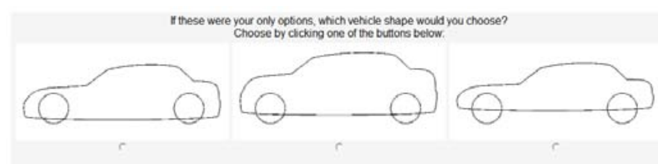

Figure 7: Screenshot from vehicle preference survey

Each survey consisted of 18 questions to build the preference model, 6 holdout questions, and a repeat question to check for consistency. The Sawtooth SSI web tool was used to administer the survey. Participants were recruited using Amazon M-turk. There were a total of 109 survey participants. The average participant age was 38.5 years old; there were 45 male and 64 female participants. Each participant was over 18.

Equations 15 and 16 describe the part-worth main and first order interaction effects models respectively:

$$v_j = \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j} + \beta_5 x_{31j} + \beta_6 x_{33j} + \beta_7 x_{41j} + \beta_8 x_{43j}, \quad (15)$$

$$v_j = \beta_1 x_{11j} + \beta_2 x_{13j} + \beta_3 x_{21j} + \beta_4 x_{23j} + \ldots + \beta_{31} x_{33j} x_{41j} + \beta_{32} x_{33j} x_{43j}, \quad (16)$$

where x_{11} and x_{13} correspond to ground clearance, x_{21} and x_{23} to body height, x_{31} and x_{33} to hood length, and x_{41} *and* x_{43} to trunk length.

In this case no information about which, if any, interactions are important is available before hand. Consequently, there is no guidance on whether a main effects model can sufficiently represent preference. As shown previously, the impact interactions can have on model performance ranges from negligible to critical. A full factorial of the 81 profiles would require 27 choice sets. This would be a fairly long survey, especially with the inclusion of holdout questions. As an alternative, each participant was given a random survey. The random approach has been shown to be robust in situations importance of interactions is not known before hand [12].

Results

Table 10: Vehicle design part-worth results

	β	Null Model	Eq. 15	Eq.16
	x_{11}	0	-0.40 (0.07) ***	-0.58 (0.19) ***
	x_{13}	0	0.03 (0.07)	0.00 (0.18)
	x_{21}	0	-0.22 (0.07) ***	-0.24 (0.18)
	x_{23}	0	0.02 (0.07)	0.54 (0.18) ***
	x_{31}	0	0.13 (0.07) *	0.48 (0.18) ***
	x_{33}	0	0.01 (0.07)	-0.48 (0.19) ***
	x_{41}	0	0.03 (0.07)	0.50 (0.18) ***
	x_{43}	0	0.02 (0.07)	-0.17 (0.19)
	$x_{11}*x_{21}$	0		0.19 (0.17)
	$x_{11}*x_{23}$	0		-0.32 (0.17) *
	$x_{11}*x_{31}$	0		-0.27 (0.18)
	$x_{11}*x_{33}$	0		0.74 (0.18) ***
	$x_{11}*x_{41}$	0		0.21 (0.17)
	$x_{11}*x_{43}$	0		0.03 (0.18)
	$x_{13}*x_{21}$	0		0.15 (0.17)
	$x_{13}*x_{23}$	0		-0.17 (0.16)
	$x_{13}*x_{31}$	0		-0.14 (0.16)
	$x_{13}*x_{33}$	0		0.37 (0.17) **
	$x_{13}*x_{41}$	0		-0.24 (0.17)
	$x_{13}*x_{43}$	0		0.17 (0.17)
	$x_{21}*x_{31}$	0		-0.72 (0.17) ***
	$x_{21}*x_{33}$	0		0.19 (0.17)
	$x_{21}*x_{41}$	0		-0.44 (0.18) **
	$x_{21}*x_{43}$	0		0.64 (0.17) ***
	$x_{23}*x_{31}$	0		-0.17 (0.17)
	$x_{23}*x_{33}$	0		0.28 (0.17)
	$x_{23}*x_{41}$	0		-0.60 (0.17) ***
	$x_{23}*x_{43}$	0		-0.55 (0.17) ***
	$x_{31}*x_{41}$	0		-0.14 (0.17)
	$x_{31}*x_{43}$	0		0.29 (0.18)
	$x_{33}*x_{41}$	0		-0.10 (0.17)
	$x_{33}*x_{43}$	0		0.03 (0.17)
In Sample	LL	-2155.48	-2121.25	-2033.80
	EAL	33%	34%	35%
	HR	30%	40%	47%
	MASE	-	-	-
Hold Out	LL	-718.49	-706.15	-680.34
	EAL	33%	34%	35%
	HR	19%	47%	45%
	MASE	16%	16%	14%

*** $p \leq 0.01$, ** $p \leq 0.05$, * $p \leq 0.1$

The Mixed Logit Estimation by Maximum Simulated Likelihood script developed by Train [35] was used to solve for the fixed β coefficients, no random coefficients are present. The part-worth model results presented in Table 10 show that both the main and interaction effects models performed better than the null model. Some of the interactions in this survey were found to be statistically significant, and the presence of interactions slightly improved predictive capability measured via MASE. The remaining error is still higher than in the simulated examples. We believe this is a result of the complex vehicle shape as well as the heterogeneity of respondent preferences for vehicle shape. Future work will examine randomized surveys with real respondents for the simpler shape (vase) as well as explore alternative formulations for representing vehicle preference shape (higher order interactions, random coefficients, etc.).

CONCLUSION

In this work the role of interaction effects in visual conjoint was explored through three examples. The first take away from this work is that the importance of interactions is dependent not only on the consumer but also on the model. Interactions are present if preference for one aspect of a shape depends on the value of another. However, the line angle example showed how model specification decisions could reduce or exaggerate the impact interactions have on a model. The vase example illustrated how there is no performance advantage associated with including interactions in a situation where consumers' preference in based solely on main effects (as modeled)..

The next issue discussed in this work is survey design. On average the D-efficient surveys performed far better than the null models. Still, we were able to find examples of surveys that were 100% efficient but resulted in vastly different performance. Modelers should keep this in mind when settling on survey designs. Using test data is one way to determine which survey design is best for a given study.

Finally, in the absence of prior knowledge of whether or not interactions are important given a model parameterization, we recommend using randomized designs to complete visual conjoint studies. This approach requires more data but is capable of estimating all interactions and averaging out any bias associated with how the design profiles are organized into choice sets.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation CAREER Award 0747911 and CMMI 1233864 and by Ford Motor Company. The views expressed are those of the authors and not necessarily those of the sponsors.

REFERENCES

[1] Green P. E., and Rao V. R., 1971, "Conjoint measurement for quantifying judgmental data," Journal of Marketing Research, **8**(3), pp. 355–363.

[2] Huber J., and Zwerina K., 1996, “The importance of utility balance in efficient choice designs,” Journal of Marketing research, **33**(3), pp. 307–317.

[3] Luo L., Kannan P. K., Besharati B., and Azarm S., 2005, “Design of Robust New Products Under Variability: Marketing Meets Design,” Journal of Product Innovation Management, **22**(2), pp. 177–192.

[4] Besharati B., Luo L., Azarm S., and Kannan P. K., 2006, “Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach,” Journal of Mechanical Design, **128**(4), pp. 884–892.

[5] Olewnik A., and Lewis K., 2006, “A decision support framework for flexible system design,” Journal of Engineering Design, **17**(1), pp. 75–97.

[6] Orsborn S., Cagan J., and Boatwright P., 2009, “Quantifying Aesthetic Form Preference in a Utility Function,” Journal of Mechanical Design, **131**(6), pp. 061001–1.

[7] Green P. E., and Srinivasan V., 1978, “Conjoint Analysis in Consumer Research: Issues and Outlook,” Journal of consumer research, **5**(2), pp. 103–123.

[8] Wu C. F. J., and Hamada M. S., 2011, Experiments: planning, analysis, and optimization, Wiley.

[9] Huber J., 2005, “Conjoint analysis: how we got here and where we are (An Update),” Sawtooth Software Conference, **98382**(360).

[10] Kuhfeld W., 2001, “Multinomial logit, discrete choice modeling An introduction to designing choice experiments, and collecting, processing and analyzing choice data with SAS.,” SAS Institute TS-643.

[11] Zwerina K., Huber J., and Kuhfeld W., 1996, “A General Method for Constructing Efficient Choice Designs,” Durham, NC: Fuqua School of Business, Duke University.

[12] Chrzan K., and Orme B., 2002, “Sawtooth Software An Overview and Comparison of Design Strategies for Choice-Based Conjoint Analysis,” **98382**(360).

[13] Sándor Z., and Wedel M., 2005, “Heterogeneous conjoint choice designs,” Journal of Marketing Research, **42**(2), pp. 210–218.

[14] Johnson R. M., 1987, “Adaptive conjoint analysis,” Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing, pp. 253–265.

[15] Swamy S., Orsborn S., Michalek J., and Cagan J., 2007, “Measurement of headlight form preference using choice-based conjoint analysis,” Proceedings of ASME IDETC/ …, (1994), pp. 1–10.

[16] Petiot J.-F., and Dagher A., 2010, “Preference-oriented form design: application to cars’ headlights,” International Journal on Interactive Design and Manufacturing (IJIDeM), **5**(1), pp. 17–27.

[17] Reid T. N., Gonzalez R. D., and Papalambros P. Y., 2010, “Quantification of Perceived Environmental Friendliness for Vehicle Silhouette Design,” Journal of Mechanical Design, **132**(10), p. 101010.

[18] MacDonald E., Lubensky A., Sohns B., and Papalambros P. Y., 2009, “Product semantics and wine portfolio optimisation,” International Journal of Product Development, **7**(1), pp. 73–98.

[19] Kelly J. C., Maheut P., Petiot J.-F., and Papalambros P. Y., 2011, “Incorporating user shape preference in engineering design optimisation,” Journal of Engineering Design, **22**(9), pp. 627–650.

[20] Tseng I., Cagan J., and Kotovsky K., 2011, “Learning Stylistic Desires and Generating Preferred Designs of Consumers Using Neural Networks and Genetic Algorithms,” Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE) 2011, Washington, DC, USA.

[21] Sylcott B., Cagan J., and Tabibnia G., 2011, “Understanding of Emotions and Reasoning During Consumer Tradeoff Between Function and Aesthetics in Product Design,” Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE) 2011, Washington, DC, USA.

[22] Turner H., Orsborn S., and Lough K. G., 2009, “Quantifying Product Color Preference in a Utility Function,” Proceedings of 2009 American Society of Engineering Management, Springfield, MO, USA.

[23] Hagerty M., 1986, “The cost of simplifying preference models,” Marketing Science, **5**(4), pp. 298–319.

[24] Green P. E., and Srinivasan V., 1990, "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," The Journal of Marketing, **54**(4), pp. 3 –19.

[25] Thurston D. L., 2001, "Real and Misconceived Limitations to Decision Based Design With Utility Analysis," Journal of Mechanical Design, **123**(2), p. 176.

[26] Abbas a. E., 2009, "Multiattribute Utility Copulas," Operations Research, **57**(6), pp. 1367–1383.

[27] Ross A. M., Hastings D. E., Warmkessel J. M., and Diller N. P., 2004, "Multi-attribute tradespace exploration as front end for effective space system design," Journal of Spacecraft and Rockets, **41**(1), pp. 20–28.

[28] Kulok M., and Lewis K., 2007, "A Method to Ensure Preference Consistency in Multi-Attribute Selection Decisions," Journal of Mechanical Design, **129**(10), p. 1002.

[29] Gensch D., and Recker W., 1979, "The multinomial, multiattribute logit choice model," Journal of Marketing Research, **16**(1), pp. 124–132.

[30] Train K. E., 2003, Discrete choice methods with simulation, Cambridge university press.

[31] Louviere J., and Woodworth G., 1983, "Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data," Journal of marketing research, **20**(4), pp. 350–367.

[32] Orme B. K., Alpert M. I., and Christensen E., 1997, "Assessing the Validity of Conjoint Analysis--Continued."

[33] Hyndman R. J., and Koehler A. B., 2006, "Another look at measures of forecast accuracy," International Journal of Forecasting, **22**(4), pp. 679–688.

[34] Kuhfeld W. F., Tobias R. D., and Garratt M., 1994, "Efficient Experimental Design with Marketing Research Applications," Journal of Marketing Research, **31**(4), p. 545.

[35] Train K., 1998, "Recreation demand models with taste differences over people," Land economics, **74**(2), pp. 230–239.