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

DETC2013-12737

EYE-TRACKING DATA PREDICTS IMPORTANCE OF PRODUCT FEATURES AND SALIENCY OF SIZE CHANGE

Ping Du
Graduate Student
Mechanical Engineering
Iowa State University
Ames, IA, USA
pdu@iastate.edu

Erin F. MacDonald
Assistant Professor
Mechanical Engineering
Iowa State University
Ames, IA, USA
erinmacd@iastate.edu

ABSTRACT

Features, or visible product attributes, are indispensable product components that influence customer evaluations of functionality, usability, symbolic impressions and other qualities. Two basic components of features are visual appearance and size. This work tests whether or not eye-tracking data can (1) predict the relative importances between features, with respect to their visual design, in overall customer preference; and (2) identify how much a feature must change in size in order to be noticeable by the viewer. The results demonstrate that feature importance is significantly correlated with a variety of gaze data. Results also show that there are significant differences in fixation time and count for noticeable vs. unnoticeable size changes. Logistic models of gaze data can predict both feature importance and saliency of size change.

1 INTRODUCTION

Product visuals are an important determinant of customer preference in almost all product categories. The preference for the overall visual design can be thought as based, in-part or in-whole, on the preferences for the visual design of individual product features. In this work, we use the term “feature” to refer to a product attribute or characteristic that is visible. Here we focus on two challenges associated with the visual design of products: determining which features are the most visually important to customers; and how much the size of a feature can change before a customer notices.

Both of these challenges, importance and size, are directly linked to the profitability of a design. Designers cannot spend equal amounts of time perfecting all visual features. They must focus on those that are most important to the customer. Likewise, production budgets for intricate molds, labor-intensive manufacturing processes, and expensive materials must be weighted towards investing in product features that are most likely to increase sales. Size concerns present budgetary constraints as well, for example a company may have the opportunity to save ten percent on production costs by reducing the size of a product feature by two percent, but may have worries that customers will notice this change and perceive it as a loss of quality or luxury. Visual appearance and size of features can both be constrained by product function and other product objectives, such as weight.

As compared to a survey approach for gathering such information, eye tracking offers more information with less exposure to stimuli. Gaze data has been used to indicate attribute importance [1], but the relationship between eye-tracking data and importance ratings has not been directly proven. This paper lays a foundation for future use of the eye-tracking data to facilitate product design.

The paper proceeds as follows: Section 2 provides background information for eye tracking research, and attribute importance and size; Section 3 contains research propositions and associated hypotheses for this paper; and Section 4 specifies the methodology. Results regarding feature importance are presented in Section 5, and results regarding size changes are provided in Section 6. Section 7 and Section 8 discuss the results and present conclusions.

2 BACKGROUND

2.1. Eye Tracking Research

Gaze data provides quantitative information on the visual acquisition of information. Eye-tracking devices collect this gaze data and corresponding software refine and analyze it for study. According to the "eye-mind" hypothesis, what people look at is an indication of what they are mentally processing [2, 3]. Gaze data provides insights into human cognitive processes to facilitate the investigation of the origins of decisions or behaviors [4] and has been used in research areas including psychology [5-7], marketing [8-10], human-computer interaction [4, 11, 12] and industrial engineering [13].

Eye-tracking has become one of the major process tracing methods for information acquisition research [14]. Another major process tracing method is computerized process tracing (CPT) which is usually conducted through the Mouselab software. The Mouselab software [15, 16] displays information on the computer with covered boxes. People acquire the information by moving the mouse curser over a particular box. In the end, the software provides details about which boxes have been visited, the time spent on each box and so forth. This kind of output is similar to that from the eye tracking process, but as Lohse and Johnson [14] identified, eye-tracking technology can monitor the process of how the information is acquired more completely and naturally.

There are a number of eye-tracking studies related to the work presented here. Pieters and Warlop [9] used eye tracking technology to study visual attention during brand choice. They found that, on average, respondents had longer fixation times on the brand they eventually chose compared with other alternatives; neither time pressure nor task motivation altered this relationship. Gofman et al. [17] found that the first gaze location on food packages was correlated with both the total amount of time spent on the packages and the purchase decisions.

Koivunen et al. [18] analyzed gaze path data to study how people perceived product designs with different given tasks: memorizing the product, evaluating its aesthetics, usability and durability. They also tested how the products were evaluated when no instructions were given. They observed that gaze paths and fixation times varied for the different tasks. Reid et al. [19] used both the eye tracking data and survey data to elucidate how customer judgments were affected by different representations of product design.

Eye-tracking equipment can be used while investigating 3D surroundings, but it is most typically used in conjunction with a computer monitor or screen, as is the case in this study. The screen presents different visual stimuli to a subject as he or she proceeds through an experimental session. Prior to data analysis, gaze data is parsed into Areas of Interest (AOIs), or areas of a given stimulus related to the research hypothesis [11]. Fixation time and fixation count are commonly-analyzed types of gaze data. "Fixations are eye movements that stabilize the retina over a stationary object of interest [20]." Fixation time refers to the duration of one fixation, and count refers to the number of fixations. Data on percentage-fixation time and first-located time are also used in this paper. The percentage-fixation time is the fixation time spent on an AOI divided by the total fixation time spent on the stimulus. The first-located time for an AOI is a measurement of the time between initial exposure to a stimulus and first fixation on that AOI. Information about additional eye tracking measurements can be found in [11, 12].

2.2. Attribute Importance

Addressing feature importance, specifically the importance of the visual design of features, can be thought of as studying a particular type of product attribute importance. Relative attribute importance identifies product attributes that are most likely to change customer preference through variation in attribute configuration. Bettman et al. define customer decision-making rules, such as compensatory and lexicographic decision rules [21] for which attribute importance can either directly or indirectly determine product choice. They model customer preference decisions (choices) by assigning different importances to different product attributes. In this model, differences in attribute importance cause each attribute of a choice option to have a weighted subjective value. These values are added together to get the total utility for the option. In this model, the customer's final choice decision largely depends on the attributes that are most important to the customer, due to the larger weights on these attributes.

Attribute importance has been assessed in different ways. The most direct way of estimating attribute importance is to ask respondents why they choose a product option. By collecting the attributes indicated in an interview, the relative importance of an attribute can be estimated by the number of times the attribute is mentioned [22]. Attribute importance can also be estimated by establishing relationships between attributes and preference decisions or other evaluations [22]. Banks [23] applied linear discriminant functions to relate preference ratings on attributes to preference of overall products; functions' coefficients were then "converted to units of the standard deviation of the corresponding variable" to indicate the relative importance of the attributes.

In conjoint or discrete choice analysis, attribute importance can be interpreted from the estimated part-worths of attribute/levels (configurations). Orsborn et al. [24] apply this specifically to visual product features. They estimated customer preferences for quantified aesthetic forms using a logit model, and mentioned that attribute importance was indicated by the magnitude of the estimated part-worths. MacDonald et al. [25] studied importance of product attributes, but not visual features. They refined the definition of attribute importance in product design, in order to perform statistical tests on this metric in a discrete choice study. Importance was defined as the percentage of customer choice that is determined by a specific attribute, in a hypothetical market where a full factorial combination of products is available. Jaccard et al. [26] conducted an information search task and gained insight on how customers searched for information about different attribute dimensions while making automobile purchasing decisions. Respondents evaluated a choice with available product profiles. Each profile had nine attribute dimensions, each with associated information available to the subjects. The authors calculated two indices of importance for each respondent: the order and number of pieces of information collected by subjects. In a mobile phone purchasing case study, Reisen et al. [27] used eye-tracking to test the relationship between attribute importance rankings and the frequency of evaluating related text. They found that the two variables were highly correlated, but there was potential bias in the non-random ordering of the related text.

There are other studies that address importance of components of alternatives, that do not study products *per se*. Jaccard and King [28] estimated attribute importance by comparing two conditional probabilities, defined as the absolute difference between the probability of an intention, such as the intention to vote for a candidate, with a presence of an attribute and without. Schkade and Johnson [29] used the Mouselab system to investigate how people evaluated two-payoff gambles in two response modes, pricing and choice, separately. They used the duration of time spent on an attribute as a measure of attention and indirectly demonstrated that the amount of attention that an attribute attracts may be an indication of its salience or importance.

2.3. Feature Size

Designers determine feature size using customer preference, technical requirements, and other sources of input. For example, a large grill on a car promotes engine cooling and better performance, but may look ugly to customers. Designers want a size change to be noticed when it has positive effects on customer preference, for example, “30% more free” in a detergent bottle. But designers work to hide, conceal, or diminish size changes that could have negative effects on customer preference, for example a decrease in car trunk size vs. last year’s model. Noticeable vs. unnoticeable difference is referred to as “saliency,” and this term is used in a binary sense (salient or not).

The relationship between attribute size and customer preference has been studied in the marketing literature. Michalek et al. [30] observed that large number size on a dial-readout scale was preferred as it indicated easy readability. Coelho do Vale et al. [31] discussed that package sizes of tempting products, small vs. large, could affect customer choices through the activation of self-regulatory. With self-regulation activated, customers were more likely to approach small packages, believed to help regulate consumption. Chandon and Ordabayeva [32] found that compared with supersizing a product in three dimensions (height, width and length), supersizing a product in only one dimension largely increased its choice share. This relationship was not affected by the fact that volume increase was clearly marked. A product downsized in three dimensions would have larger choice share than that downsized in one dimension. These results were due to the visual bias that same amount of volume change through three dimensions was considered smaller than that through one dimension. Yang and Raghubir [33] have found that elongated containers for frequently purchased goods were considered to have a larger volume which could lead to decreased purchase quantity. Krider et al. [34], studied the perception of container shape and showed that a rectangular cream cheese container was considered larger than a round one even though they actually had the same volume, leading people to buy a lower quantity if packaged as a rectangle. Krider et al. also discovered that customers initially relied on a single dimension, which was most salient, to make comparisons between areas.

3 RESEARCH PROPOSITIONS AND ASSOCIATED HYPOTHESES

Proposition 1: Feature importance is correlated with gaze data in preference choices between two products. This proposition is inferred from and supported by the literature presented in Section 2.1 and 2.2, see [1, 9, 17, 22, 26, 29]. The proposition is tested by the following hypotheses. This first set of hypotheses is accompanied by explanations in plain English to assist in understanding.

- *Hypothesis 1a: There is a positive correlation between feature importance and the feature’s fixation time.* It is hypothesized that subjects spend a longer time looking at more important features during the choice task, and that the longer they look, the more important the feature.
- *Hypothesis 1b: There is a positive correlation between feature importance and the feature’s percentage-fixation time.* It is hypothesized that subjects spend a larger percentage of time looking at important features than other features.
- *Hypothesis 1c: There is a positive correlation between feature importance and the feature’s fixation count.* It is hypothesized that subjects look more frequently at important features than other features.
- *Hypothesis 1d: There is a negative correlation between feature importance and the feature’s first located time.* It is hypothesized that subjects look at important features first.

Proposition 2: Saliency of size change can be predicted by gaze data. Sütterlin et al. [1] used gaze data to examine how customers evaluated pairs of options, which were described by text information and were shown sequentially. Some information provided in a pair was the same between options while the other information in the two options was different. In this way, the two options in a pair had both shared and unique information. They observed that the shared information between the two options was evaluated normally when it appeared in the first option but almost ignored when it appeared again in the second option. Based on their findings, we expect that this phenomenon will appear when the size change of a feature for a pair is unnoticeable. Gaze data, therefore, could be used to detect the saliency of size changes by testing whether features are ignored or not.

People use different viewing strategies for evaluating stimuli sequentially vs. side-by-side, so the two conditions are tested with different hypotheses. "Product B" refers to the right-hand-side product in the side-by-side condition and the second product seen in the sequential condition. Two more measurements, Δ fixation time and Δ fixation count, are defined to test the proposition. Δ fixation time/count represents the difference in these quantities for features appearing on Product B vs. Product A. The proposition is tested by a number of hypotheses summarized on the next page and in Table 1. Note that the blank cells in Table 1 are also tested in the analysis, for completeness.

- *Hypothesis 2a: A noticeable-size-change of a feature in Product B has a longer fixation time than an unnoticeable one (sequential).*
- *Hypothesis 2b: A noticeable-size-change of a feature in Product B has a higher fixation count than an unnoticeable one (sequential).*
- *Hypothesis 3a: A noticeable-size-change feature pair has a longer total fixation time than an unnoticeable one (side-by-side).*
- *Hypothesis 3b: A noticeable-size-change feature pair has a higher total fixation count than an unnoticeable one (side-by-side).*
- *Hypothesis 4a: Δ fixation time of a noticeable-size-change feature is different from that of an unnoticeable-size-change feature (sequential).*
- *Hypothesis 4b: Δ fixation count of a noticeable-size-change feature is different from that of an unnoticeable-size-change feature (sequential).*

4 METHODOLOGY

To test the hypotheses, a computer-based survey was designed and implemented in Attention Tool software by iMotions [35]. Two products, cars and electric bicycles, were used, described further in Section 4.1. Participants were instructed to sit in front of a Tobii T120 Eye-tracker and completed the entire survey through its 17 inch TFT monitor with a 1024 x 768 pixel resolution. The survey started automatically after the participant completed a calibration process. Instructions were provided ahead of each section. Participants had to complete a practice question in each experimental section before seeing the real test products and the associated questions. Questions for each stimulus were presented after the stimulus on separate pages.

Table 2 provides an overview of the survey which will be described in detail in Section 4.2. Different sections served different goals of the study. Section I and Section II, associated with the information collected in Section IV, tested Proposition 1 with its associated Hypotheses: 1a-1d. Section III tested Proposition 2 with its associated Hypotheses: 2a, 2b, 3a, 3b, 4a and 4b. Product images for Section I of the survey had only variable features (no size changes). Images used in Section II, and Section III had size manipulations and variable features.

4.1. Stimuli

Sample stimuli used in the survey are shown in Fig.1. The 2012 Chevy Cruze from the Chevrolet website [36] and a Shanyang electric bicycle model from the Global-tradekey website [37] were used as basic models for the stimuli. Only one image for each of the products was used, different perspectives of the products were not shown. This car was selected because it provided a basic sedan model that was familiar to customers. This electric bicycle was selected because it was transformed from a bike model, which made for a product that was familiar in some ways, but unfamiliar in others. As compared to a car, visually processing the bicycle required more effort.

Sets of stimuli were created from these base images in Adobe Photoshop. The car stimuli included four visual design variants each for the headlights, grill, side mirrors, and wheels. Design variants were also created for the handlebars, footrest, seat and cargo box of the electric bicycle. These design variants are numbered and can be found in Table A.1 in the Appendix. To create these variants for the two products, features of different cars and electric bicycles were "pasted" onto the base images and carefully blended into the images. A pilot study was conducted to evaluate response to the images. Subjects found that the images looked natural and did not distract them from evaluating the product designs. In other words, the design variant components were not noticeably different from the rest of the image. The feature used to generate design variations will be referred to as a *"varied feature"* in the rest of the paper.

For size manipulations, feature designs from the base images were proportionally resized as discussed below. They were created for the headlights, grill, and side mirrors of the car and the handlebars, seat, and cargo box of the electric bicycle. Each of these six features will be referred to as a "*target feature*" in the rest of the paper when discussing size changes. Three pairs of products were generated for each target feature to create three levels of size differences for the feature. Due to proportionalities in the base images, the levels of size differences for cars were 15%, 20%, and 25% while those for electric bicycles were 10%, 15%, 20%. For each pair of target features, Product A displayed the base size of the feature and Product B displayed the manipulated size. The designs of the remaining varied features were randomly picked from their corresponding design pools with the restriction that none of them share the same design in a pair. A text illustration of the three pairs generated for size manipulations of headlights is shown in Table 3 as an example. Numbers in the table stand for different design variants of the varied features (see Table A.1 in the Appendix), and the percentages represent enlargements.

TABLE 1. An illustration of the test for Proposition 2 about size-changes.

		Condition	
AOI	**Fixation metric**	Sequential	Side-by-side
Size-changed feature in Product B	Time	H2a: Noticeable, longer	
	Count	H2b: Noticeable, higher	
Size-changed feature in Product A and B combined	Time		H3a: Noticeable, longer
	Count		H3b: Noticeable, higher
Size-changed feature in Product B vs. Product A	Delta time	H4a: Noticeable, longer	
	Delta count	H4b: Noticeable, higher	

TABLE 2. An overview of the survey design.

Sect.	Survey Questions	Images	Conditions	Hypotheses	Gaze Data Used	Survey Data Used
I	1) Indicate preferences 2) Rate satisfaction of (1) 3) Rate Product A 4) Rate Product B	Feature Design Variants	Sequential & Side-by-Side	1a	Fixation Time	No
				1b	%-fixation Time	
				1c	Fixation Count	
				1d	First-located Time	
II	Indicate preferences	Size Manipulation and Feature Variants	Sequential & Side-by-Side	1a	Fixation Time	No
				1b	%-fixation Time	
				1c	Fixation Count	
				1d	First-located Time	
III	Identify and write down the features that are different between two products	Size Manipulation and Feature Variants	Sequential	2a	Fixation Time	Features (with size manipulation) mentioned by subjects were classified as "noticeable-size-change"; Features (with size manipulation) that were not mentioned were classified as "unnoticeable-size-change"
				2b	Fixation Count	
				4a	Δ fixation Time	
				4b	Δ fixation Count	
			Side-by-Side	3a	Fixation Time	
				3b	Fixation Count	
IV	Rate importance for different features	None	None	1a, 1b, 1c, 1d	None	Ratings used to test correlations
V	Demographic questions	None	None	None	None	No

4.2. Survey Design

The survey had five sections, as summarized in Table 2, which took about twenty minutes complete. Sample questions asked in the survey are provided in Table A.2 in the Appendix. Section I showed images as sequential (ISeq) or side-by-side (ISBS). In ISeq, Product A is shown first; in ISBS, Product A is shown on the left. Subjects in both conditions evaluated two randomly-assigned pairs of cars and then two pairs of electric bicycles. After evaluating each pair of products, participants were instructed to complete four tasks: (1) indicate preferences using an 8-level scale, which ranged from "Strongly prefer Product A" to "Strongly prefer Product B"; (2) rate their satisfaction with the preference decision they just made using an 8-level scale, ranging from "Very unsatisfied" to "Very satisfied"; (3) rate Product A using an 8-level scale, ranging from "Very bad" to "Very good"; (4) rate Product B using an 8-level scale, ranging from "Very bad" to "Very good." These scales are from Houston and Sherman [38]. Using 8-level scales in these tasks forces participants to either prefer Product A or B, for example. This data is being analyzed for cancelation/focus behavior [38] in related work.

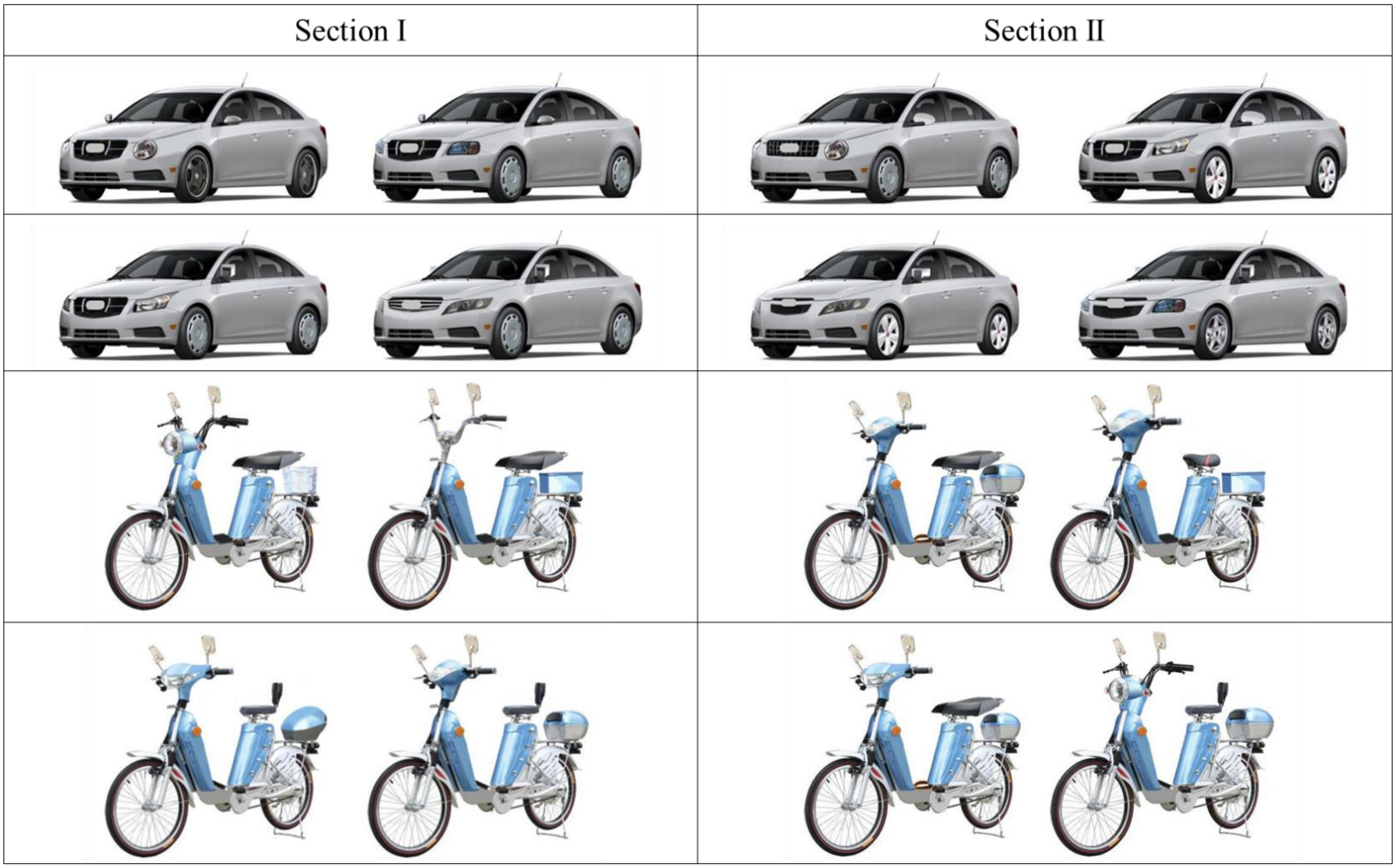

FIGURE 1. Sample pairs used in the survey (size manipulations in pairs from Section II are Side mirror (20%), Grill (25%), Handlebar (20%), and Cargo box (20%), from top to bottom).

TABLE 3. Car designs for size manipulations of headlights.

		Headlight	Grill	Wheel	Side Mirror
Pair 1	Car A	1	4	2	1
	Car B	1 at 115%	1	1	4
Pair 2	Car A	1	3	3	2
	Car B	1 at 120%	2	4	3
Pair 3	Car A	1	1	2	3
	Car B	1 at 125%	2	1	1

Section II also showed images as sequential (IISeq) or side-by-side (IISBS). Subjects in both conditions evaluated three randomly-assigned pairs of cars and then three pairs of electric bicycles, with the requirement that in each product category the participant was shown each level of size change once and each target feature with size manipulation once. As an example, three pairs of cars presented to a participant could be: a pair in which the headlights were manipulated for a size change at the 15% level; a pair in which the grill was manipulated for a size change at the 25% level; and a pair in which the side mirrors were manipulated for a size change at the 20% level. In this survey section, participants were asked to indicate their preferences for each pair of products using 8-level scales, which ranged from “Strongly prefer Product A” to “Strongly prefer Product B”.

Section III also showed images as sequential (IIISeq) or side-by-side (IIISBS). For each participant, stimuli presented in Section III were repeated from Section II. The participant was asked to identify and write down the features that were different between the two options within a pair, rather than indicating preferences. This allowed for comparison in size noticing between different types of tasks: implicit and explicit size evaluations, which will be addressed in future work.

Section IV collected the importance ratings for all features of the car and of the electric bicycle, including those not varied in the experiment. Table A.3 in the Appendix contains a full list of features mentioned in the survey. Participants were instructed to consider the importance of each feature in forming their preferences towards the product and then rate the importance on 7-level scales, in which level 1 stood for “Not important at all”, and level 7 stood for “Very important”.

The survey ended with Section V which collected some demographic information about the participants.

4.3. Participants

The survey was designed for 72 participants. However, due to initial computer issues which resulted in unrecorded data for 11 participants, a total of 83 adults from Iowa State University took the survey, compensated with $5 or extra credit. An online screening survey was conducted to make sure that the potential participants met the eye-tracking study requirements suggested by Pernice and Nielsen [39]. Participants had normal to corrected vision; did not wear bifocals, trifocals, layered lenses, or regression lenses; did not have difficulty reading a computer screen unassisted; and did not have cataracts, eye implants, glaucoma or permanently dilated pupils [19, 39]. Table 4 provides a summary of the demographic information of the participants. Participants were randomly assigned to either sequential or side-by-side conditions.

TABLE 4. A summary of subject demographic information.

Gender		Age	
Category	Number of Participants	Category	Number of Participants
Male	37	18-24	27
		25-34	22
		35-44	15
Female	35	45-54	4
		55-64	3
		65-74	1

4.4. Data Preparation

The gaze data, together with the survey responses, were collected and managed by the Attention Tool software. An Area of Interest (AOI, see Section 2.1), was manually created for each feature of the product following its original outline, as shown in Fig. 2. After AOIs were generated for each stimulus, gaze data was exported to R, a free statics software platform, for post-processing, and then analyzed using the statistical software package JMP.

While the iMotions software worked well to track eye movements and identify fixations overall, it failed for two participants, which had very few fixations identified for almost all product stimuli. An additional nine participants had no fixations for only a few stimuli. These missing fixations were identified in post-processing and treated as missing data in the analysis.

FIGURE 2. An example of the AOIs generated for a car.

5 PROPOSITION 1 RESULTS

The relationships between importance ratings and gaze data were tested based on the gaze data from Section I and Section II of the survey separately. The gaze data from Section II was used to fit an ordinal logistic model to predict the importance of product features.

5.1. Survey Section I Results

Subject- or individual-level averages, in terms of the fixation time, the percentage-fixation time, the fixation count and the first-located time, were calculated for each feature (including both varied and unvaried features). As one subject saw two pairs of products for each product category, four measurements (of fixation time, %time, count, and first-located time) were averaged for each feature. Then, Pearson correlations between each type of individual-level gaze data for a feature and corresponding individual-level importance rating were tested for the ISeq and the ISBS conditions.

Test results are shown in Table 5. In both the ISeq and ISBS conditions, there are significant positive correlations between feature importance rating (from Section IV) and fixation time, percentage-time and count. There is a significant negative correlation between feature importance rating and the first-located time on the feature. Hypotheses 1a-1d are strongly supported by these results.

To visualize change in gaze data across feature importance levels, observations of how a feature was evaluated by each participant were grouped into seven sets based on the subject's importance rating. This process was performed for the fixation time, percent time, and count separately. Similar trends were obtained for each, so fixation time is used as a demonstration. As shown in Fig. 3, for both the ISeq condition and the ISBS condition, there is a clear trend showing that the average fixation time spent on the feature is longer when its importance level is higher. These trends are consistent with the significant positive correlations found for Hypothesis 1a. The average first-located time was plotted with the importance level, shown in Fig. 4. Again, the first-located time for a feature decreases with its importance level in ISeq; while in the ISBS condition, the pattern is less clear. This may be due to the limited data in that condition (12 participants).

TABLE 5. Correlations between feature importance and associated gaze data (subject-level averages) for Sections I and II.
('*' $p<0.05$, '**' $p<0.01$, '***' $p<0.0001$)

Fixation Metric	Time (ms)	%Time	Count	First-located (ms)
ISeq	0.24**	0.33***	0.26**	-0.24**
ISBS	0.22**	0.26***	0.25**	-0.16*
IISeq	0.21***	0.25***	0.23***	-0.23***
IISBS	0.25***	0.27***	0.25***	-0.17***

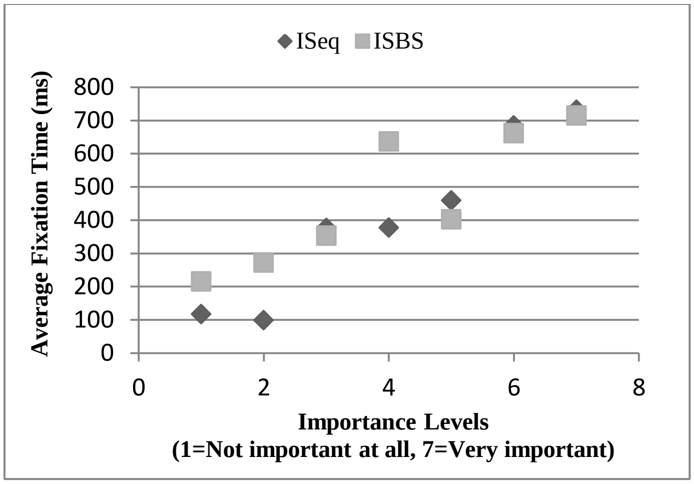

FIGURE 3. Average fixation time spent on a feature increases with its importance level (Section I).

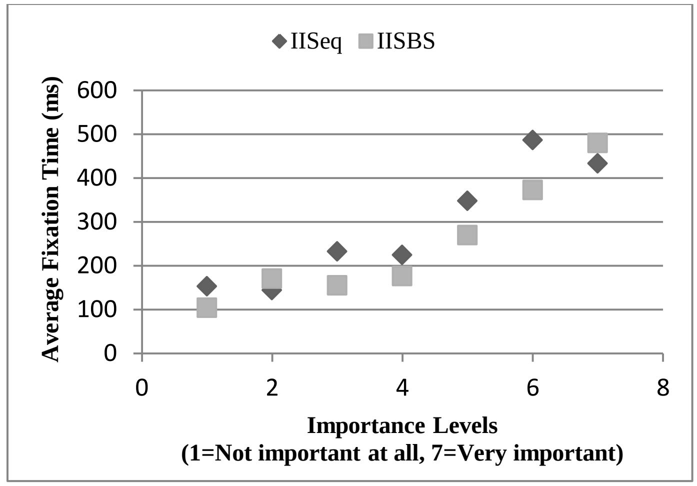

FIGURE 5. Average fixation time spent on a feature increases with its importance level (Section II).

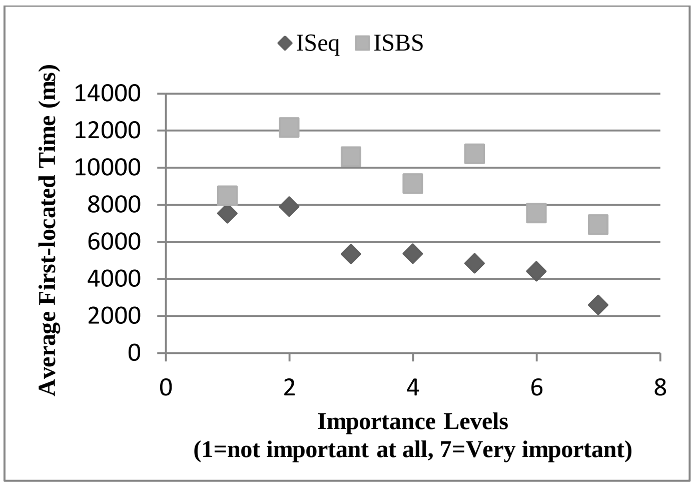

FIGURE 4. Average first-located time on a feature varies with importance levels (Section I).

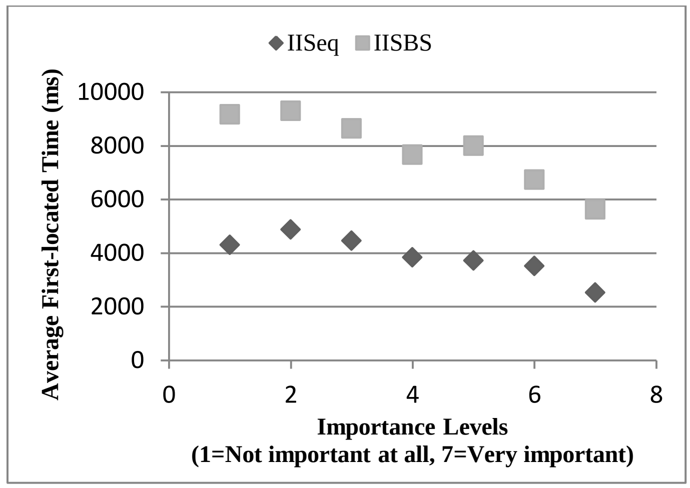

FIGURE 6. Average first-located time on a feature decreases with its importance level (Section II).

5.2. Survey Section II Results

Correlations. The method described in Section 5.1 was used here to test the relationships between feature importance and gaze data. Results of the Pearson correlations are shown in Table 5. In both conditions, there are significantly positive correlations between feature importance rating and the fixation time, percent time and count. There is a significantly negative correlation between feature importance rating and the first-located time. Hypotheses 1a, 1b, 1c, and 1d are strongly supported by these results. Graphs of correlations are shown in Figs. 5 and 6. In both conditions, there is a clear trend showing that the average fixation time spent on a feature increases and the average first-located time on the feature decreases as its importance level goes up. This indicates that features considered as more important are examined earlier and for longer. These trends are consistent with the significant correlations that are demonstrated in Table 5.

Ordinal Logistic Model. An ordinal logistic regression was applied to predict feature importance level using gaze data. This regression can appropriately model the ordinal trends of the importance levels which ranged from “Not important at all” to “Very important”. To avoid any potential multicollinearity problems [40], caused by involving correlated variables, only one type of gaze data should be chosen as the independent variable. Here, the percentage-fixation time was chosen for the reason that it had the largest correlation with the feature importance rating in both the IISeq and IISBS conditions. Two models were fit, one for the Sequential condition and the other one for the Side-by-Side condition. The summary of the fit is shown in Table 6. It indicates good fit of the models for both conditions. The percentage fixation time significantly affects both models, indicating its correlation with the feature importance.

TABLE 6. The ordinal logistic models for predicting importance level using % fixation time. Each intercept represents an importance level (plus one reference level).

	IISeq		IISBS	
Prob>ChiSq	<.0001		<.0001	
Term	**Estimate**	**Std Error**	**Estimate**	**Std Error**
Intercept[1]	-3.54***	0.27	-3.47***	0.26
Intercept[2]	-2.08***	0.15	-2.29***	0.16
Intercept[3]	-0.47***	0.10	-0.56***	0.10
Intercept[4]	-0.16	0.10	-0.27**	0.09
Intercept[5]	1.44***	0.11	1.26***	0.11
Intercept[6]	2.66***	0.15	2.84***	0.16
%Fix. Time	-0.07***	0.01	-0.17***	0.02

6 PROPOSITION 2 RESULTS

Based on the write-in responses from survey Section III (indicating what size changes are noticed), data from the AOIs of target features (size-manipulated features) were classified into two sets. One set included noticeable-size-changes and the other set included unnoticeable-size-changes. These two sets were compared using fixation time/count and Δ fixation time/count. The results from the IIISeq and IIISBS conditions were analyzed separately and tested by One-way ANOVA.

6.1. Results from Survey Part III, Sequential (IIISeq)

As detailed in Table 7, noticeable-size-changes in Product B have significantly larger values of average fixation time and count than unnoticeable ones. When considering the AOIs of the target feature in Product A and Product B combined, noticeable-size-changes have a significantly larger average fixation time and count than unnoticeable ones. Average Δ fixation time for the noticeable-size-changes is significantly different from the unnoticeable ones, with the former value above zero and the latter one below zero. Average Δ fixation count shows the similar results. These results strongly support Hypotheses 2a, 2b and 4a, 4b.

A nominal logistic regression was applied to predict saliency of a size change using gaze data. Here, fixation count for target feature on Product B was chosen as the independent variable for the model. When fitting by this fixation metric, the model had the largest Chi-Square value, indicating best fit, compared with those fit by other available metrics. A summary of the fit is provided in Table 8. It demonstrates good fit of the model and also validates that fixation count for a target feature on Product B has a significant effect on differentiating noticeable and unnoticeable size changes.

TABLE 8. The nominal logistic model for predicting saliency of size changes (with the unnoticeable size change as reference level) for Sequential condition.

Prob>ChiSq	<.0001	
Term	**Estimate**	**Std Error**
Intercept	2.40***	0.29
Fix. Count for target feature on Product B	-0.29***	0.07

6.2. Results from Survey Part III, Side-by-Side (IIISBS)

The analysis in Section 6.1 was performed for data from the IIISBS condition. Detailed results are demonstrated in Table 7. Noticeable-size-changes in Product B have significantly larger values of average fixation time and count than the unnoticeable-size-change ones. When the AOIs of the target feature in Product A and B are combined, noticeable-size-changes have significantly larger values of average fixation time and count than unnoticeable ones. Average Δ fixation time and count show no differences between the two sets of features that are compared. Hypotheses 3a and 3b are strongly supported here.

A nominal logistic regression was applied to predict saliency of a size change using gaze data for Side-by-Side condition. Total fixation time spent on the paired target features was used to fit the model because it provided the best fit among all the available fixation metrics. A summary of the fit is provided in Table 9. It indicates good fit of the model and also validates that fixation time spent on the paired target features has a significant effect on differentiating noticeable and unnoticeable size changes.

TABLE 9. The nominal logistic model for predicting saliency of size changes (with the unnoticeable size change as reference level) for Side-by-Side condition.

Prob>ChiSq	<0.001	
Term	**Estimate**	**Std Error**
Intercept	1.56***	0.23
Fix. time for paired target feature	-0.0004**	0.0001

7 DISCUSSION

The survey results support both research propositions: (1) feature importance is correlated with gaze data in preference choices between two products, and (2) Saliency of size changes can be predicted by gaze data.

TABLE 7. A summary of results for Proposition 2

AOI	Fix. metric	Condition	
		Sequential	Side-by-side
		Noticeable vs. Unnoticeable	Noticeable vs. Unnoticeable
Size-changed feature in Product B	Time	1080ms vs. 400ms ***	906ms vs. 483ms **
	Count	4.13 vs. 1.63 ***	3.66 vs. 2.38 **
Size-changed feature in Product A and B combined	Time	1843ms vs. 1110ms *	1697ms vs. 919ms ***
	Count	7.22 vs. 4.33 **	7.25 vs. 4.43 **
Size-changed feature in Product B vs. Product A	Delta time	319ms vs. -303ms **	115ms vs. 47ms
	Delta count	1.03 vs. -1.06 **	0.07 vs. 0.32

Hypotheses 1a-1d hold true in all situations tested here. During the processes of making preference decisions, there are positive correlations between feature importance and three types of gaze data: the fixation time, the percentage-time and the count. Each shows a clear trend with increasing feature importance. These findings are consistent with Bettman et al.'s conclusions: people pay more attention to the information that has a larger weight in achieving the decision goal [21]. There is a negative correlation between feature importance and the feature's first-located time. As the feature importance rises, it's first-located time decreases. The results from the Side-by-Side condition in Section I did not show a clear decreasing trend. As mentioned in Section 5.1, this may be due to the limited data in that condition. In Section I, each condition has only 12 participants while in Section II, each condition has 36 participants. In Section I, each participant only evaluated two pairs for a product category while in Section II each participant evaluated three pairs for a product category. Evaluating more pairs is more likely to indicate the true measurement of how the participant evaluates a feature.

Results from Section III of the survey show that the saliency of a size change can be predicted with gaze data. In both the IIISeq and IIISBS conditions, noticeable and unnoticeable size changes can be differentiated with gaze data. Hypotheses 2a and 2b are strongly supported by the data of IIISeq. A noticeable-size-change in Product B has significantly larger values of fixation time and count than an unnoticeable one. These findings are extended through the test of Δ fixation time and count. These tests prove that, within a noticeable-size-change feature pair, a target feature in Product B has longer fixation time and higher fixation count than that in Product A; while within an unnoticeable-size-change feature pair, the target feature in Product B has shorter fixation time and lower fixation count than that in Product A. Considering the fixation time and count for the target feature in Product A as a reference, the fixation time and count in Product B behave oppositely for the noticeable and unnoticeable size changes. The noticeable-size-change attracts extra attention; while the unnoticeable one is ignored. These trends can be used to explain the results in Hypotheses 2a and 2b. Therefore, when a feature with an unnoticeable size change appears in two products shown sequentially, its latter appearance (in Product B) is considered as a repetition of the former one and is less competitive in attracting attention. These findings are consistent with [1, 38]: when a pair of products is evaluated for preference decisions, their shared information is likely to be ignored in its second appearance. Even though not originally hypothesized, results show that when AOIs of a target feature in Product A and B are combined, a noticeable-size-change has significantly longer total fixation time as well as higher a count. These results validate the findings in Hypotheses 4a and 4b.

In IIISeq, the break between the two stimuli weakens the memory of the first stimulus. It is possible that only abstract representations of the first stimulus remain while its details are overwritten by the second stimulus, according to the "overwriting" explanation for "change blindness" [41]. Even in this situation, there are some size changes that trump overwriting and are salient with the viewer. This is worth further investigation, as it is unlikely that a customer will, in the real world, compare two slightly different product variations directly to each other.

As Hypotheses 3a and 3b predict, in the IIISBS condition, a noticeable-size-change feature pair has significantly larger fixation time and count than an unnoticeable one. This indicates that subjects are paying less attention to the unnoticeable-size-change feature pair. Even not though hypothesized, in the IIISBS condition, a noticeable-size-change in Product B has significant larger values of fixation time and count than an unnoticeable one. This may have resulted from the fact that gaze typically moves from left to right. The Δ fixation time and count do not support a hypothesis for the IIISBS condition. This suggests that different viewing strategies are adopted in the sequential and side-by-side conditions. Presenting stimuli side by side enables pairwise comparisons between options, so it is less likely that the saliency of a size change will be identified by comparing attention assigned to the target features in Product B vs. A.

All hypotheses are tested under two general conditions, showing product images sequentially and side-by-side. Based on the results, the relationships found between feature importance and the gaze data are almost the same in the two conditions. This suggests that researchers studying the importance of product attributes using eye tracking can present two product images at a time (such as in choice decisions) and reliably draw conclusions about attribute importance. It is not necessary to show product images individually. But the two conditions have different results when the Δ fixation time and count are used to differentiate noticeable and unnoticeable size changes: the Δ fixation time and count are useful only in the Sequential condition.

8 CONCLUSION

Results from this study show that feature importance is correlated with a variety of gaze data (fixation time, percentage-time, and count; and first-located time). Gaze data is validated as providing insight into how people evaluate the features while making preference decisions. The importance level of a product feature can be predicted by the gaze data using the ordinal logistic model. This study also proves that not all the size changes of features can be identified. Gaze data provides clear evidence that differentiates noticeable and unnoticeable size changes.

This study could be furthered by developing a method to determine the just-noticeable threshold for size changes, which would be immediately useful to practicing designers. The different gaze data associated with implicit and explicit size evaluation tasks could be compared. The study can also be further developed by investigating the effect of product viewing perspective, sizes, etc. on the results. The method of identifying saliency of size changes in this paper is applicable to identify saliency of manufacturing imperfections or geometrical variations. A considerable amount of time and money has been spent on manufacturing processes to ensure the quality appearance of products [42, 43]. If one can predict how likely it is that an imperfection or variation will be noticed, optimization analysis can be performed to reduce the manufacturing costs while maintaining the targeted quality appearance of products. These findings help designers make decisions by providing a new approach to identify the importance and saliency of product features. They suggest that feature importance can be identified at the individual-person-level in only three questions, significantly reducing a subject’s mental burden associated with other methods such as using discrete choice analysis or complex rating procedures. The efficiency of the approach suggests a range of applications including product personalization.

ACKNOWLEDGMENTS

The authors would like to thank the iMotions company for providing the Attention Tool software and the associated technical support; Dr. Frederick Lorenz, Dr. Heike Hofmann, and Kent Kroeger at Iowa State University for statistical advice; and. colleagues in the IRIS Design Lab in Iowa State University for their valuable inputs.

REFERENCES

[1] Sütterlin, B., Brunner, T. A., and Opwis, K., 2008, "Eye-Tracking the Cancellation and Focus Model for Preference Judgments," Journal of Experimental Social Psychology, **44**(3), pp. 904–911.

[2] Just, M. A., and Carpenter, P. A., 1976, "Eye Fixations and Cognitive Processes," Cognitive Psychology, **8**(4), pp. 441–480.

[3] Nielsen, J., and Pernice, K., 2010, *Eyetracking Web Usability*, New Riders, Berkeley, CA.

[4] Schiessl, M., Duda, S., Tholke, A., and Fischer, R., 2003, "Eye Tracking and Its Application in Usability and Media Research," MMI-interactiv Journal, **6**(6), pp. 1-10.

[5] Rayner, K., 1998, "Eye Movements in Reading and Information Processing: 20 Years of Research," Psychological Bulletin, **124**(3), pp. 372-422.

[6] Findlay, J. M., and Gilchrist, I. D., 1998, "Eye Guidance and Visual Search," *Eye Guidance in Reading and Scene Perception*, G. Underwood, eds., Elsevier, Oxford.

[7] Glaholt, M. G., and Reingold, E. M., 2011, "Eye Movement Monitoring as a Process Tracing Methodology in Decision Making Research," Journal of Neuroscience, Psychology, and Economics, **4**(2), pp. 125-146.

[8] Rosbergen, E., Wedel, M., and Pieters, R., 1990, "Analyzing Visual Attention to Repeated Print Advertising Using Scanpath Theory," Research Institute SOM, University of Groningen, Groningen, Netherlands.

[9] Pieters, R., and Warlop, L., 1999, "Visual Attention During Brand Choice: The Impact of Time Pressure and Task Motivation," International Journal of Research in Marketing, **16**(1), pp. 1-16.

[10] Pieters, R., and Wedel, M., 2004, "Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-size Effects," Journal of Marketing, **68**(2), pp. 36-50.

[11] Jacob, R. J. K., and Karn, K. S., 2003, "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises," *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, J. Hyona, R. Radach, and H. Deubel, eds., pp. 573-605.

[12] Poole, A., and Ball, L. J., 2005, "Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects," *Encyclopedia of human computer interaction*, C. Ghaoui, eds., Idea Group Reference, Pennsylvania.

[13] Konstantopoulos, P., 2009, "Investigating Drivers' Visual Search Strategies: Towards an Efficient Training Intervention," Doctor of Philosophy, University of Nottingham, Nottingham, UK.

[14] Lohse, G. L., and Johnson, E. J., 1996, "A Comparison of Two Process Tracing Methods for Choice Tasks," Organizational Behavior and Human Decision Processes, **68**(1), pp. 28-43.

[15] Payne, J. W., Bettman, J. R., and Johnson, E. J., 1988, "Adaptive Strategy Selection in Decision Making," Journal of Experimental Psychology: Learning, Memory, and Cognition, **14**(3), pp. 534-552.

[16] Dhar, R., Nowlis, S. M., and Sherman, S. J., 1999, "Comparison Effects on Preference Construction," Journal of Consumer Research, **26**(3), pp. 293-306.
[17] Gofman, A., Moskowitz, H. R., Fyrbjork, J., Moskowitz, D., and Mets, T., 2009, "Extending Rule Developing Experimentation to Perception of Food Packages with Eye Tracking," Open Food Science Journal, **3**, pp. 66-78.
[18] Koivunen, K., Kukkonen, S., Lahtinen, S., Rantala, H., and Sharmin, S., 2004, "Towards Deeper Understanding of How People Perceive Design in Products," *CADE2004 Web Proceedings of Computers in Art and Design Education Conference*, M. A. Eriksen., L. Malmborg, and J. Nielsen, eds., Sweden.
[19] Reid, T., MacDonald, E., and Du, P., "Impact of Product Design Representation on Customer Judgment with Associated Eye Gaze Patterns," *Proceedings of the ASME 2012 International Design Engineering Technical conferences & Computers and Information in Engineering conference*, Chicago, IL, August 12-15.
[20] Duchowski, A. T., 2007, *Eye Tracking Methodology: Theory and Practice*, Springer, London.
[21] Bettman, J. R., Luce, M. F., and Payne, J. W., 1998, "Constructive Consumer Choice Processes," Journal of Consumer Research, **25**(3), pp. 187-217.
[22] Myers, J. H., and Alpert, M. I., 1968, "Determinant Buying Attitudes - Meaning and Measurement," Journal of Marketing, **32**(4), pp. 13-20.
[23] Banks, S., 1950, "The Relationship Between Preference and Purchase of Brands," Journal of Marketing, **15**(2), pp. 145-157.
[24] Orsborn, S., Cagan, J., and Boatwright, P., 2009, "Quantifying Aesthetic Form Preference in a Utility Function," Journal of Mechanical Design, **131**(6), p. 061001.
[25] MacDonald, E. F., Gonzalez, R., and Papalambros, P., 2009, "The Construction of Preferences for Crux and Sentinel Product Attributes," Journal of Engineering Design, **20**(6), pp. 609-626.
[26] Jaccard, J., Brinberg, D., and Ackerman, L., 1986, "Assessing Attribute Importance: A Comparison of Six Methods," Journal of Consumer Research, **12**(4), pp. 463-468.
[27] Reisen, N., Hoffrage, U., and Mast, F. W., 2008, "Identifying Decision Strategies in A Consumer Choice Situation," Judgment and Decision Making, **3**(8), pp. 641-658.
[28] Jaccard, J., and King, G. W., 1977, "The Relation Between Behavioral Intentions and Beliefs: A Probabilistic Model," Human Communication Research, **3**(4), pp. 326-334.
[29] Schkade, D., and Johnson, E., 1989, "Cognitive Processes in Preference Reversals," Organizational Behavior and Human Decision Processes, **44**(2), pp. 203-231.
[30] Michalek, J. J., Ebbes, P., Adiguzel, F., Feinberg, F. M., and Papalambros, P. Y., 2011, "Enhancing Marketing with Engineering: Optimal Heterogeneous Markets," International Journal of Research in Marketing, **28**(1), pp. 1-12.
[31] Coelho do Vale, R., Pieters, R., and Zeelenberg, M., 2008, "Flying under the Radar: Perverse Package Size Effects on Consumption Self-Regulation," Journal of Consumer Research, **35**(3), pp. 380-390.
[32] Chandon, P., and Ordabayeva, N., 2008, "Downsize in 3D, Supersize in 1D: Effects of the Dimensionality of Package and Portion Size Changes on Size Estimations, Consumption, and Quantity Discount Expectations," INSEAD Working Papers Collection, **46**.
[33] Yang, S., and Raghubir, P., 2005, "Can Bottles Speak Volumes? The Effect of Package Shape on How Much to Buy," Journal of Retailing, **81**(4), pp. 86-96.
[34] Krider, R. E., Raghubir, P., and Krishna, A., 2001, "Pizzas: π or Square? Psychophysical Biases in Area Comparisons," Marketing Science, **20**(4), pp. 405-425.
[35] iMotions, http://imotionsglobal.com/, Accessed 4/21/2013.
[36] Chevrolet, http://www.chevrolet.com/flash.html, Accessed 4/21/2013.
[37] Global-tradekey, http://www.global-tradekey.com/company/G02189/C64201.htm, Accessed 4/21/2013.
[38] Houston, D. A., and Sherman, S. J., 1995, "Cancellation and Focus - The Role of Shared and Unique Features in The Choice Process," Journal of Experimental Social Psychology, **31**(4), pp. 357-378.
[39] Pernice, K., and Nielsen, J., 2009, "Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking," Nielsen Norman Group, Fremont, CA.
[40] Paul, R. K., "Multicollinearity: Causes, Effects and Remedies."
[41] Simons, D. J., 2000, "Current Approaches to Change Blindness," Visual Cognition, **7**(1/2/3), pp. 1-15.
[42] Forslund, K., Kero, T., and Söderberg, R., 2010, "Appearance Fmea - A Method for Appearance Quality Evaluation of Early Design Concepts," *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, **8**, pp. 217-225.
[43] Söderberg, R., Wickman, C., and Lindkvist, L., 2008, "Improving decision making by simulating and visualizing geometrical variation in non-rigid assemblies," Cirp Annals-Manufacturing Technology, **57**(1), pp. 175-178.

APPENDIX

TABLE A.1. Design pool for varied features.

Car					Electric Bicycle				
Design Variants	Headlight	Grill	Wheel	Side mirror	Design Variants	Handlebar	Seat	Cargo box	Footrest
1					1				
2					2				
3					3				
4					4				

TABLE A.2. Sample questions asked in the survey about cars.

Section I	1. Please compare Car B to Car A and indicate your preference using the following scale. 2. Please evaluate your decision according to the following requirements. (1) Please think about the car you prefer in this pair and rate your satisfaction with the decision using the following scale. (2) Please rate Car A using the following scale. (3) Please rate Car B using the following scale.
Section II	Given these two options of cars, which one do you prefer? Please use the following scale to rate your preference for the two cars.
Section III	Please identify the differences between these two cars (just list the names of the parts which are different).
Section IV	We have divided a basic car model into nine components, hood/windshield, grill, headlight, bumper/lower grill, wheel, side door, side mirror, side window, and tail as shown in the image below. How important are these different components' design in forming your preference for the car? Please rate the importance for the design of these components respectively using the following scales.

TABLE A.3. A collection of features rated in the survey.

Car	Hood/ Windshield	Grill	Headlight	Bumper/ Lower grill	Wheel	Side mirror	Side window	Side door	Tail	
Electric Bicycle	Rearview mirror	Handlebar	Front frame	Tire	Footrest	Pedal	Rear frame	Seat	Cargo box	Kick stand