ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
IDETC2016
August 21-24, 2016, Charlotte, North Carolina

DETC2016-60293

EVALUATIONS THAT MATTER: CUSTOMER PREFERENCES USING INDUSTRY-BASED EVALUATIONS AND EYE-GAZE DATA

Priya Seshadri[1]
Graduate Student
Mechanical Engineering
Purdue University
West Lafayette, IN, United States

Youyi Bi
Graduate Student
Mechanical Engineering
Purdue University
West Lafayette, IN, United States

Jaykishan Bhatia
Undergraduate Student
Mechanical Engineering
Purdue University
West Lafayette, IN, United States

Ross Simons
Global Product Research Analyst
General Motors
Warren, MI, United States

Jeffrey Hartley
Technical Director
General Motors
Warren, MI, United States

Tahira Reid
Assistant Professor
Mechanical Engineering
Purdue University
West Lafayette, IN, United States

ABSTRACT
This study is the first stage of a research program aimed at understanding differences in how people process 2D and 3D automotive stimuli, using psychophysiological tools such as galvanic skin response (GSR), eye tracking, electroencephalography (EEG), and facial expressions coding, along with respondent ratings. The current study uses just one measure, eye tracking, and one stimulus format, 2D realistic renderings of vehicles, to reveal where people expect to find information about brand and other industry-relevant topics, such as sportiness. The eye-gaze data showed differences in the percentage of fixation time that people spent on different views of cars while evaluating the "Brand" and the degree to which they looked "Sporty/Conservative", "Calm/Exciting", and "Basic/Luxurious". The results of this work can give designers insights on where they can invest their design efforts when considering brand and styling cues.

Keywords: Eye-tracking, Product Evaluation, Automotive Design, Vehicle Design, 2D, 3D

1 INTRODUCTION

The design of an automobile exterior on the road is the end product of a long iterative process. As a design is developed, its depiction evolves from just a written description to 2D depictions (first rough sketch ideas and then realistic drawings) to 3D models (first scale models but ultimately full size models). At each stage, decisions are made about which designs should go forward and how to evolve them. An important factor in many of these decisions is feedback from customers. It is therefore important that design decision makers have faith in the feedback received from customers.

It is only natural that as a byproduct of this process, the decision makers form intuitions about how promising each design is. When feedback from customers runs counter to these expectations, it is also natural for the decision makers to question the results, and the stimulus format is often an easy target for blame.

The two stimuli formats each have pros and cons. For 3D exteriors, results are more believable for decision makers and surface contours are more apparent, since respondents have infinite viewing angles. Furthermore, the respondent is able to see the actual size and proportions. But 3D properties take longer to develop, at a much higher cost and they cannot be shared globally without expensive and time-consuming shipment. Furthermore, because of the time needed to build a high quality property, each theme is a more dated "snapshot" due to development time for the property. The sheer cost and resources needed to build a 3D property also result in fewer themes being tested.

In contrast, 2D stimuli are cheaper and faster to develop and modify, facilitating the testing and iteration of more themes. They are easier to control, enabling researchers to reduce demand characteristics. And they can be moved electronically, enabling testing at multiple locations at no extra cost. But decision makers have concerns about results from 2D studies, because respondents are not seeing the real object. Respondents must imagine the 2D representation of the vehicle as a full-size object and the image may be somewhat ambiguous due to the limited number of angles available for view. This could arguably lead to confusion about surface contours and the true shape of the exterior, especially as designs become more dramatically styled.

The ideal process would use 2D for early research, enabling the testing and rapid iteration of a variety of designs, and then when the candidates have been winnowed down to just one design, perhaps switching to a 3D model. But this is contingent on assuring decision makers that they can trust 2D results.

To address these concerns, GM has conducted many studies that show the conclusions would be the same or very similar, regardless of stimulus format. This research-on-research usually has consisted of running parallel studies, using the same designs with matched samples. The only difference between these studies has been the stimulus format (2D vs. 3D). Results have consistently shown a strong positive correlation between the results of these matched studies, and company leadership would have made the same decisions on theme selection regardless of stimuli format. Despite these findings, decision makers still resist the use of 2D stimuli.

The current research program will complement this previous work by exploring the <u>process</u> of perceiving and judging 2D and 3D stimuli. By "process" we mean, as people are judging stimuli, we will monitor psychophysiological measures such as their level of arousal and emotion and where they are looking. Previous studies have just looked at the <u>product</u> of these judgments (i.e., the ratings) and took similar rating patterns as support for substituting one format for the other. The current research program will be assessing perceptual and psychophysiological activity as it happens. One claim made by automotive designers is that 2D depictions do not inspire as emotional a reaction as does a 3D model. The current research program will ultimately address that claim where this paper serves as first step.

2 BACKGROUND LITERATURE

In this section, we review the existing literature on the impact of different representation modes in product evaluation. We show the research gap within existing measures to capture customers' responses in product evaluation: few studies include more measures beyond self-report. The necessity to utilize biometric signals for monitoring individuals' evaluation process is discussed. We believe that eye-tracking is a promising tool to provide insights on how product representations can deliver specific meanings to customers. Bridging this gap would allow us to further understand how customers perceive the visual information of a product while evaluating associated attributes.

2.1 Research on the impact of shape and representation modes in product evaluation

In user-centered design, customers' responses, e.g. subjective preferences and judgments to products (especially for shapes), have been utilized extensively to assess the value of product concepts (Orsborn et al. [1]). Orbay, Fu & Kara [2] investigated the relationship between product form and consumer responses through a visual deconstruction and abstraction of existing final products. They produced a spectrum of abstractions for a given 3D computer model (e.g. a 9-level simple-to-complex abstraction of a Mustang model). Their results show that emotional responses evoked by coarse product impressions are strongly correlated with those evoked by final production models. This in turn, highlights the importance of early aesthetic assessment and exploration before committing to detail design efforts. A study by Dahan and Srinivasan [3] comparing different forms of visual depiction indicated that similar results were achieved by using virtual prototypes or physical prototypes. The authors state that certain kind of product characteristics like smell, taste and touch that may affect choices made by consumers can be communicated well through physical prototypes.

In design related studies, researchers usually present participants with products for visual evaluation, including asking for their preferences, examining purchasing decisions (e.g. Discrete Choice Analysis (Wassenaar and Chen [4]), emotions (e.g. Kansei Engineering (Nagamachi [5]) and/or other higher-level cognitive reactions to products as well, such as the perceived environmental friendliness on car silhouettes (Reid et al. [6]). Customers' shape preferences can be captured either quantitatively or semantically. For example, Swamy et al. [7] represented customer preferences for the shape of automobile headlights in a utility function directly related to the engineering representation of the headlight. A choice-based conjoint survey was used to discover and design the most preferred shape. Their study allows quantitative measurement of consumer preferences for a broad class of organic shapes and optimization of the design to maximize desirability. Yumer et al. [8] proposed a shape editing method where the user creates geometric deformations using a set of semantic attributes instead of doing detailed geometric manipulations. This method provides a platform for quick design explorations and allows non-experts to produce semantically guided shape variations. In summary, designers are able to design better products by incorporating customers' shape preferences.

In early stage design, product concepts are usually displayed with sketches and low fidelity prototypes. While in later stages, sophisticated engineering drawings, 3D CAD models and high-fidelity prototypes are often used. Researchers have noticed that different representation modes can affect how people judge product concepts. For example, Sauer and Sonderegger [9] examined the impact of prototype fidelity on user behavior, their subjective evaluations, and emotion in the context of mobile phone design. They found that participants considered computer-based prototypes (medium fidelity) to be even more attractive than fully operational appliances (high fidelity). This is a counter-intuitive result, but it reveals that individuals appeared to compensate for deficiencies in aesthetic design by overrating the aesthetic qualities of reduced fidelity prototypes. Bates-Brkljac [10] investigated whether computer-generated architectural representations are perceived as a more credible means of communicating design than traditional forms of representations. Results show that computer-generated photomontage is perceived as the most credible and perspective drawing as the least credible form of representation. Artacho-Ramirez et al. investigated how people perceived a computer loudspeaker in four different ways of graphically representing them (Artacho-Ramirez et al. [11]). Their results show that photographic representation suffices to communicate most of the concepts in the same way that the real product would.

2.2 Research on the application of eye-tracking and other biometric devices in product evaluation

Product evaluation is an information processing and decision making process, which can be monitored by many process tracing methods, such as survey/self-report (Nisbett and Wilson [12]), computer-based information board paradigms (e.g. Mouselab (Payne et al. [13])) and think-aloud protocols (Montgomery and Svenson [14]). However, these techniques sometimes influence evaluation behavior (Billings and Marcus [15]) and might hinder participants from relying on automatic processing by constraining quick comparisons and information search (Glockner and Betsch [16]). In contrast, biometric signals, such as eye gaze data, provide tracing information without hindering automatic information acquisition processes. These signals are promising to provide insights on understanding individuals' decision processes (Glockner and Herbold [17]).

Eye-tracking research is based on Just and Carpenter's eye-mind hypothesis (Just and Carpenter [18]) that people look at what they are thinking about. Accordingly, people fixate on a specific area of a product diagram longer when they are attracted or are confused. The main metrics used in eye-tracking include: (1) fixations: eye movements that stabilize the retina over a stationary object of interest (in this study, a fixation is defined as an eye-sight focus on a specific area for more than 100 milliseconds); (2) fixation time: a measure of the duration of the fixation; and (3) scan paths: connections between consecutive fixations (Goldberg and Kotval [19]). The location and duration of fixations is directly related to the locus and difficulty of cognitive processing (Loftus and Mackworth [20]). Thus, tracking eye movements may provide insight into what visual information is being processed currently and how difficult this information is to process, which may serve as an additional measure for people's thinking process (Madsen [21]).

In consumer and user experience studies, eye-tracking has been used to reveal the visual attention patterns between different groups of consumers or users. Djamasbi et al. [22] found that when looking at e-commerce websites, baby boomers had significantly more fixations and their fixations covered more of the pages (e.g., headers, main body, sidebars) than those of millennials. Meissner and Decker [23] showed that eye-tracking data can be used to verify the choice-based conjoint (CBC) results, and to learn on the individual level what attribute levels of the products might be unacceptable for the respondent. Researchers in the design community also started to apply the eye-tracking method to observe how people perceive the shapes of products. For example, Reid et al. [24] conducted a study in which subjects were shown computer sketches, front/side view silhouettes, simplified renderings, and realistic renderings to test the extent to which a variety of judgments including opinions, objective evaluations, and inferences are affected by form presentation. Results show that while inferences were consistent across form, opinions were not. Associated eye-tracking data, such as fixation times and fixation counts, provide additional insight into these findings. Du and MacDonald [25] also tested whether or not eye-tracking data can predict the importance of product features in overall customer preference. Their results indicate that feature importance is significantly correlated with a variety of gaze data, i.e. people would pay more attention to those with more important product features. See more eye-tracking studies in consumer research [26-28], product design [29, 30] and a general review [31]. These studies show the potential of applying eye-tracking in product evaluation, which provides researchers more comprehensive ways to understand individuals' decision behavior.

2.3 Summary and Hypotheses

In Section 2.1, the studies discussed showed that in most cases, computer-based digital representations (e.g., computer-based prototypes, computer generated photomontages, photographs) are able to convey product concepts or ideas to customers as expected. Section 2.2 showed that eye tracking is a promising tool to provide additional insights on the behavioral and psychophysiological processes occurring during product evaluation tasks.

This paper presents some initial work as part of a larger research program to deeply examine the effects of 2D vs. 3D during product evaluations. In this study, we are interested in how customers perceive brand information and make style judgments from three basic views of vehicle renderings – Front (F), Side (S), and Rear (R). In addition to survey questions, we use eye-tracking to better understand which views customers attend to the most during product evaluations. These insights help designers identify the most effective areas to put design cues when trying to convey certain information to customers (brand, style, etc.). We focus on reporting the eye-tracking results which help us understand if certain representations are assessed longer by participants while making evaluations.

In the following section, the details of the design of experiment and experimental procedures are presented.

3 METHODOLOGY

3.1 Design of Experiment

Figure 1 represents the general structure of our experiment. The main study consists of a computer-based survey, in which we present participants with visual stimuli of four different cars and ask them six types of rating questions about each car.

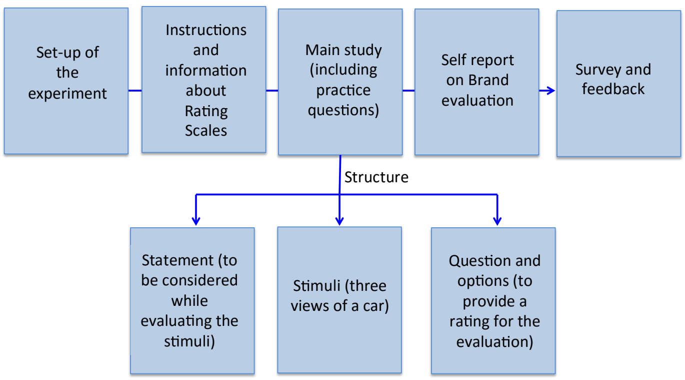


Figure 1. The overall structure of the experiment

Each of the different car stimuli used in this study consisted of three different views. The first was the front three-quarter view, which shows the front view and a partial view of the side, the second was a side view, which only shows the side of the respective car and the third one was the rear three-quarter view, which shows the rear view and a partial side view. In order to randomize the order in which the stimuli were presented in the experiment (to eliminate any effects due to the order of presentation), all three images were shown with the same size and resolution. Their position with every question was changed so as not to give one view preference over the other. Results from a pilot study were used to determine the position of the images and the question with respect to each other. Various arrangements with question at the top, middle and bottom and various arrangements of the images were tested before deciding the layout presented in Fig. 2. To control for extraneous variables, all stimuli were in the same format (renderings) and color. Brand cues and logos were masked so that the participants' answers were not based on brand loyalty. General Motors (GM) provided these images as they are typically used in their internal studies. Figure 2 illustrates an example of the visual stimulus of a car that was used in the experiment.

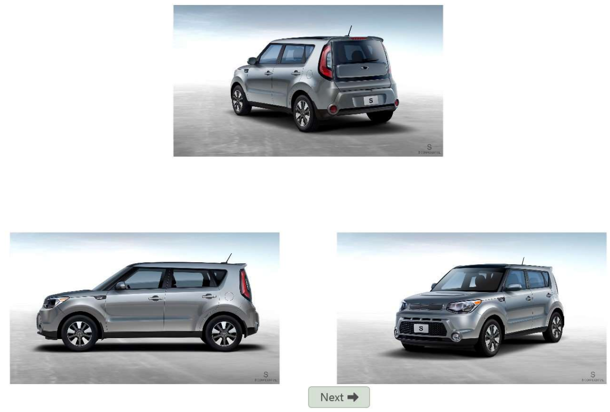


Figure 2. An example of the visual stimulus from the experiment.

The three views of the car stimulus were randomly ordered when shown to the participants (e.g. in Fig. 2, rear view is on the top, side view is at bottom left and front view is at bottom right; while for other stimuli, front view might have been at the top or the side view). Six types of evaluation questions were asked when participants perceived the visual stimuli of these cars:

a) Please evaluate the overall appeal of the car.
b) Please evaluate if the car is Sporty or Conservative.
c) Please evaluate if the car is Traditional or Innovative.
d) Please evaluate if the car is Basic or Luxurious.
e) Please evaluate if the car is Calm or Exciting.
f) Please indicate the brand of the car.

These questions are of direct interest to the automotive industry in product evaluation. Questions b-e are semantic differentials which have a well-established history in psychology as tools for unearthing the meaning of, in this case, vehicles (Coates [33], Osgood et al. [34]). Questions a-e were evaluated by rating on a scale of 1 to 5. For example, participants were required to evaluate the stimuli for the statement "Overall exterior appearance is appealing" and recorded their response on a scale of 1 to 5, 1 being 'Disagree strongly' and 5 being 'Agree strongly'. While evaluating the other questions like Sporty-Conservative, participants were required to indicate these evaluations on a scale of 1 to 5, 1 being Sporty (or "the word Sporty applies completely to this car") and 5 being Conservative (or "the word Conservative applies completely to this car"). For question f, 30 common car brands were provided to the participants and they chose the brand that they felt the car stimulus belonged with.

The following hypotheses guided this research:

H1a: When evaluating the brand of a vehicle, there will be a difference in the times spent looking at each of the 3 views

H2a: When evaluating the degree to which a vehicle is sporty/ conservative, there will be a difference in the times spent looking at each of the 3 views

H3: For each of the other evaluations, there will be differences in the time spent looking at each of the 3 views

Insights from industry experts and a study on headlight shapes [7] help to inform the expected outcomes of these hypotheses:

H1b: Customers will look at the front view of the vehicle during brand evaluations more than the other views.

H2b: Customers will look at the side view of the vehicle during sporty/conservative more than the other views.

3.2 Participants

A total of 19 students and staff members from a large Midwest university participated in this preliminary study. There were 13 female and 6 male participants ranging in age from 19 to 60 years old (mean: 29.05 years, standard deviation: 11.12 years). Participants were recruited by email and flyers. None of them knew the stimuli or the purpose of the study and were compensated $10 for their participation. They also had to meet the inclusion criteria suggested by Pernice and Nielsen [32] to satisfy the experimental conditions of eye tracking research:

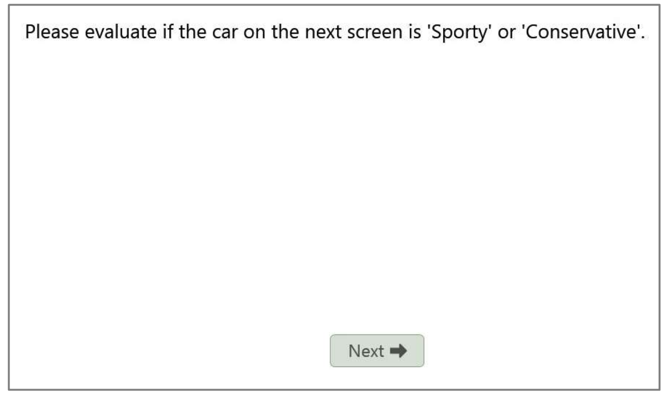
- Have normal to corrected vision (contact lenses and glasses are okay except for bifocals, trifocals, layered lenses or regression lenses).
- Do not have glaucoma, cataracts, eye implants, or permanently dilated pupils.
- Can read a computer screen and the Web without difficulty.
- Do not need a screen reader, screen magnifier or other assistive technology to use the computer and the Web.

3.3 Experimental Procedure and Data Collection

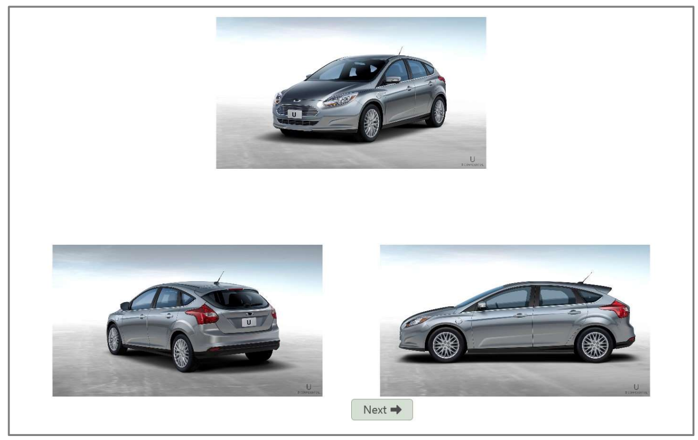
In this experiment, a Tobii X-60 (Tobii Technology AB, Danderyd, Sweden) eye-tracking device was used to capture the eye movements of the participants as they answered the survey questions. The collection rate of this eye tracker is 60 Hz. The iMotions software (iMotions Inc., Cambridge, MA) was used to present the survey and integrate and synchronize the eye-tracker with other measures including facial expression recognition and GSR.

After passing pre-screening, qualified participants were introduced to the purpose and procedures of the study by the researcher. Participants were then required to sit in front of a computer display and adjust their sitting position to ensure the successful calibration of the eye-tracker. After calibration, participants saw a welcome screen with a brief introduction to the experiment and a few practice questions before answering the real questions.

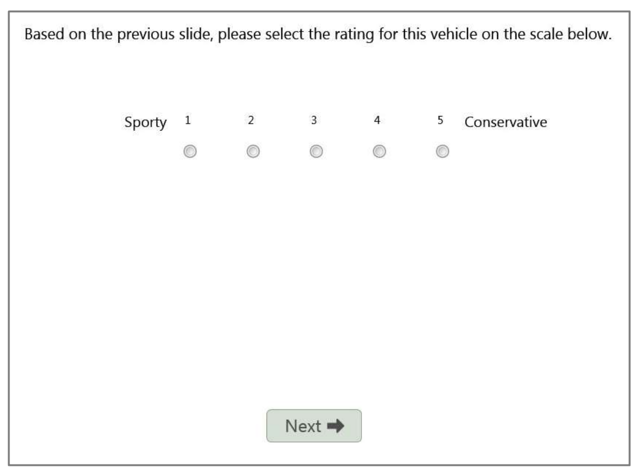
Each evaluation question had a set of three slides. Within each set, a participant first saw a slide of the question statement (Fig. 3(a)). The statement provided information that should be considered while making the evaluation. Then she/he saw a visual stimulus (which included 3 views of a car) (Fig. 3(b)). The participant made an evaluation on what they just perceived. The last slide in the set provided options so that the participant could record their response (Fig. 3(c)).



(a) Statement of an evaluation question



(b) Visual stimulus of a car design



(c) An evaluation question and options

Figure 3. The three slides of a set for an evaluation question.

Since there were four different cars and six types of evaluation questions, participants were required to repeat the above process 24 times for different questions and different stimuli. The participants also completed a survey at the end of the experiment which included questions related to their past experience and interest in cars. The participants took part in the study individually. The entire experiment lasted approximately 15-25 minutes.

The data collected in the experiment included the participants' responses while answering the car evaluation questions, their eye gaze data in specific Areas Of Interest (AOIs) of the presented visual stimuli and their responses in the post survey and self-report. We used the percentage of fixation time as the eye gaze data metric in this paper. The percentage of fixation time is calculated by dividing the fixation time on an AOI with the fixation time on the whole visual stimulus (e.g. Fig. 3 (b)). Thus larger percentage fixation time within an AOI indicates that the viewers spend more visual attention on that area.

4 ANALYSIS AND RESULTS

The eye-gaze data were analyzed using a combination of descriptive and inferential statistics. The percentage of fixation time has been reported for each car view and for each question. Percentage of fixation time represents the proportion of time that the participant spent looking at a particular AOI. The results are presented for each question. **F** represents Front view of the car, **R** represents Rear view of the car and **S** represents Side view of the car. A one-way ANOVA was conducted to compare results between each of the views to determine whether or not there were statistical differences between them. A Tukey analysis was conducted on those in which statistical differences were found. A 95% confidence level was used in both sets of analyses.

4.1 Heat-Maps of Eye-tracking Results

Heat-maps provide qualitative aggregate eye-gaze data for all participants in a given study. Figure 4(a)-4(d) shows examples of heat maps that were generated by participants while making "Brand", "Sporty/Conservative", "Calm/Exciting", and "Basic/Luxurious" evaluations of the cars, respectively. Green shading indicates lower intensity and Red shading indicates higher intensity of fixation.

Figure 4(a) shows that the Front view had the highest intensity of fixation compared to the other views while Brand was being evaluated. In other words, it received the most attention while conducting Brand evaluations.

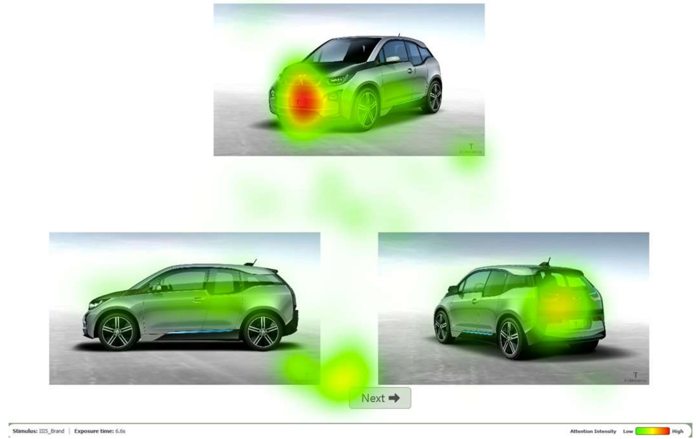


Figure 4(a): Heat map (average for all participants) of a car while the Brand is being evaluated.

Figure 4(b) shows the heat map for a car while being viewed for "Sporty/Conservative" semantic evaluations. From a qualitative standpoint, the Rear and Side view had higher intensities of fixation as compared to the Front view, indicating participants focused more on the Rear and Side views for this car.

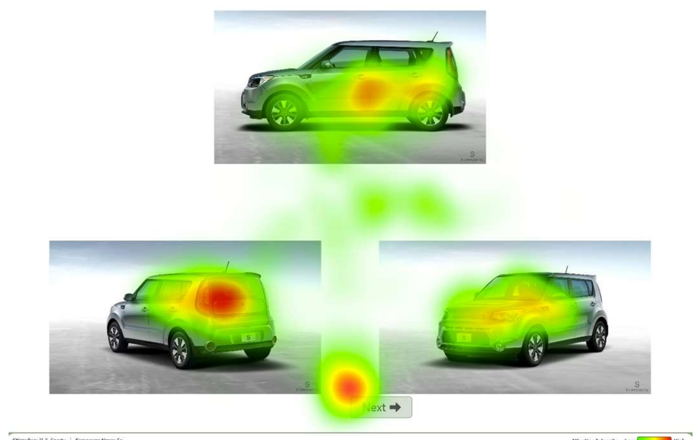


Figure 4(b): Heat map (average for all participants) of a car while it is being evaluated for being Sporty or Conservative.

Figure 4(c) shows the heat map of a car while "Calm/Exciting" evaluations were done. The Front and Side views have higher intensities of fixation indicating more focus on these views for this evaluation.

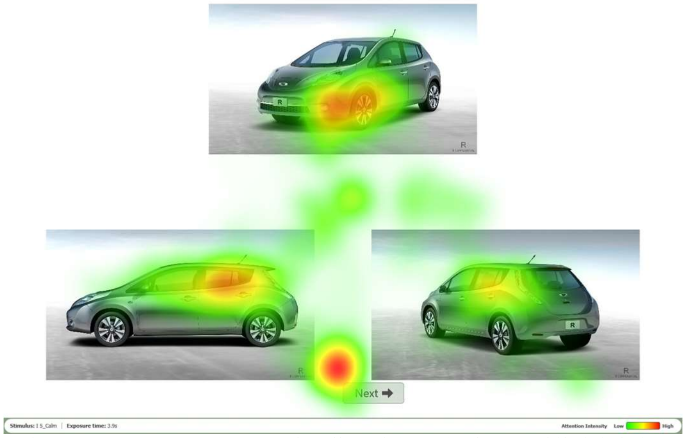


Figure 4(c): Heat map (average for all participants) of a car while it is being evaluated for being Calm or Exciting.

Figure 4(d) shows that the Side view had the highest intensity of fixation while the car was being evaluated for "Basic/Luxurious". The figure also shows that the Front view received more attention than the Rear view.

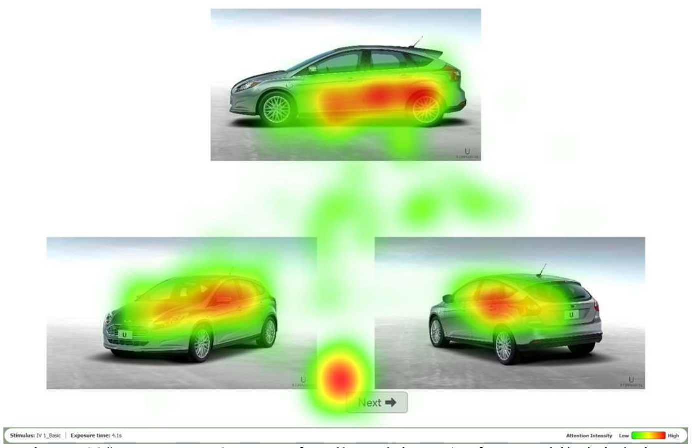


Figure 4(d): Heat map (average for all participants) of a car while it is being evaluated for being Basic or Luxurious.

4.2 Statistical Analysis of Eye-tracking Results

A one-way ANOVA was conducted on each of the six questions discussed in section 3 in order to test the five hypotheses that were proposed.

4.2.1 Results for Hypothesis 1a, 2a and 3

The results of the one-way ANOVA indicate that there was a statistically significant difference between the average percentage of fixation time for the three views while evaluating "Brand" (F(2,225) = 27.21, p = 0.000), "Sporty/Conservative" (F(2,225) = 2.98, p = 0.053), "Basic/Luxurious" (F(2,225)=9.90, p = 0.000) and "Calm/Exciting" (F(2,225)= 11.20, p = 0.000). . A Tukey post-hoc test was conducted and the results are discussed in the next section. These analyses support hypotheses 1a and 2a. There was no statistical difference between the average percentage of fixation time for the three views while evaluating "Appeal" (F(2,225) = 2.63, p = 0.075) and "Traditional/Innovative" (F(2,225)=1.65, p = 0.194). Hence hypothesis 3 is not supported by this analysis.

Figures 6-9 provide a visual representation of the percentages of fixation times for the three views while evaluating "Sporty/Conservative", "Basic/Luxurious", "Calm/Exciting", and "Brand" and show that there were differences in the time spent looking at each of the three views. The Tukey analysis in the next section will provide insights on which views they were.

4.2.2 Results for Hypothesis 1b, 2b and 3

A Tukey post-hoc test was conducted on "Sporty/Conservative", "Basic/Luxurious", "Calm/Exciting", and "Brand" Evaluations.

Table 1 below provides a summary of the Tukey results:

Table 1: Summary of the Tukey Post-hoc test results

	P-value	Tukey test?	F-R	F-S	R-S	Relevant Hypothesis
Sporty-Conservative	0.053	✓	x	x	Significant	*2a, 2b*
Basic-Luxurious	0.000	✓	x	Significant	Significant	*3*
Calm-Exciting	0.000	✓	Significant	x	Significant	*3*
Traditional-Innovative	0.194	X	---	---	---	*3*
Brand	0.000	✓	Significant	Significant	Significant	*1a, 1b*
Appeal	0.075	X	---	---	---	*3*

1) 'Brand' : F-R, F-S, R-S Significantly Different

A one-way ANOVA and the Tukey post hoc comparisons show that the average percentage of fixation time spent while evaluating "Brand" was significantly different between all three views (F(2,225) = 27.21, p = 0.000). A Tukey post-hoc test revealed that the average percentage of fixation time for the Front view (M=28.6; SD=16.3) is significantly higher as compared to that for the Rear view (M=19.1, SD=12.2; p=0) and for the Side view (M=13, SD=10.1; p=0). This number was also statistically significant when the Rear view and Side view were compared (p=0.013). This shows that hypothesis 1b is not rejected. Figure 6 provides a visual representation of the percentage of fixation time and the Standard Error associated with each value.

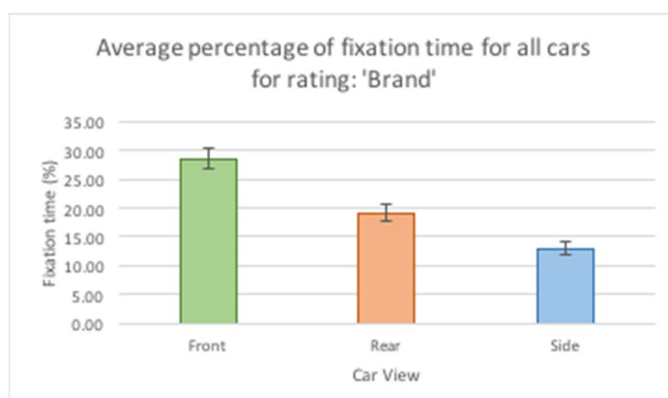


Figure 6: Average percentage of fixation time for all cars while evaluating 'Brand'

2) 'Sporty or Conservative': R-S Significantly Different

A one-way ANOVA and the Tukey post hoc comparisons show that the average percentage of fixation time spent while evaluating "Sporty or Conservative" was significantly different between all three views ($F(2,225) = 2.98$, $p = 0.053$). A Tukey post-hoc test revealed that the average percentage of fixation time for the Rear view (M=15, SD=8.7) is significantly higher as compared to that for the Side view (M=19.6; SD=10.5; p=0.048). However, the difference between the average percentage of fixation times were not significant when the Front view (M=17.1; SD=9.2) was compared to the Rear view (M=15, SD=8.7; p=0.764) and the Side view (M=19.6; SD=10.5; p=0.218). Hence hypothesis 2b cannot be completely accepted. Figure 7 provides a visual representation of the percentage of fixation time and the Standard Error associated with each value.

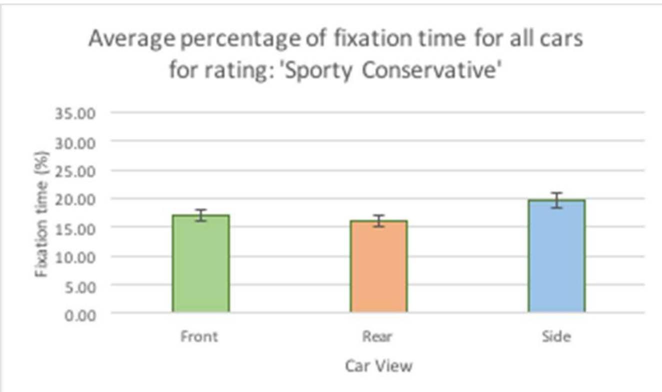


Figure 7: Average percentage of fixation time for all cars while rating 'Sporty Conservative'.

3) 'Basic or Luxurious': R-S, F-S Significantly Different

A one-way ANOVA and the Tukey post hoc comparisons show that the average percentage of fixation time spent while evaluating "Basic or Luxurious" was significantly different between all three views ($F(2,225) = 9.90$, $p = 0.000$). A Tukey post-hoc test revealed that the average percentage of fixation time for the Side view (M=20.6; SD=11.9) is significantly higher as compared to that for the Front view (M=16.3; SD=9.1; p=0.025) and for the Rear view (M=13.5, SD=8.4; p=0.000). However, the difference between the average percentage of fixation times was not significant when the Front view (M=16.3; SD=9.1) was compared to the Rear view (M=13.5, SD=8.4; p=0.171). Figure 8 provides a visual representation of the percentage of fixation time and the Standard Error associated with each value.

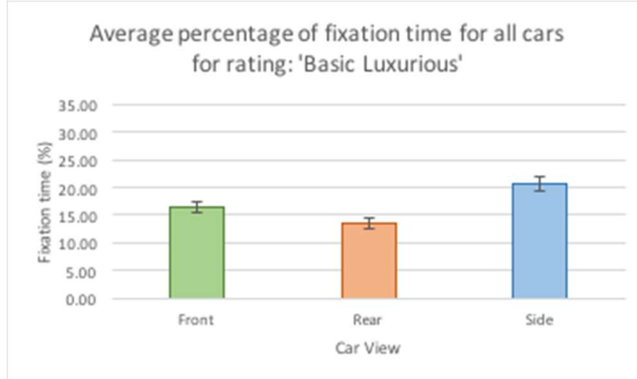


Figure 8: Average percentage of fixation time for all cars while rating 'Basic Luxurious'.

4) 'Calm or Exciting': R-S, F-R Significantly Different

A one-way ANOVA and the Tukey post hoc comparisons show that the average percentage of fixation time spent while evaluating "Calm or Exciting" was significantly different between all three views ($F(2,225) = 11.20$, $p = 0.000$). A Tukey post-hoc test revealed that the average percentage of fixation time for the Rear view (M=11.5, SD=9.3) is significantly lower as compared to that for the Front view (M=16.9; SD=9.3; p=0.001) and for the Side view (M=17.9; SD=9.4; p=0.000). However, the difference between the average percentage of fixation times was not significant when the Front view (M=16.3; SD=9.1) was compared to the Side view (M=17.9; SD=9.4; p=0.76). Figure 9 provides a visual representation of the percentage of fixation time and the Standard Error associated with each value.

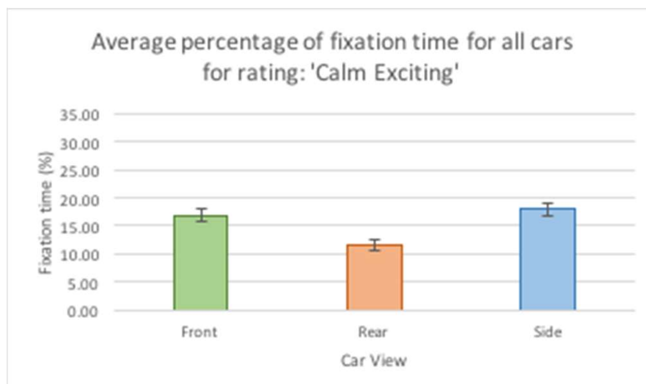


Figure 9: Average percentage of fixation time for all cars while rating 'Calm Exciting'.

There was no significant difference between the average percentages of fixation time for the three views while evaluating 'Appeal' and rating 'Traditional or Innovative', so they have not been analyzed and discussed. A summary of these results is also presented in Table 2.

Table 2: Summary of average percentage of fixation time for each question

	FRONT	SIDE	REAR
APPEAL			
SPORTY	17.1	19.6	15
INNOVATIVE			
LUXURIOUS	16.3	20.6	13.5
EXCITING	16.9	17.9	11.5
BRAND	28.6	13	19.1
OVERALL STUDY	19	18	15.4

4.2.3 Summary of Fixation Times Within the Study

Overall, for the entire study, the participants spent the most time on the Front View of the car and the least time on the Rear view of the car (Table 3). A one way ANOVA indicates that at least one of the differences is significant ($F(2,1365) = 14.28$, $p = 0.000$). The post hoc comparisons showed that the average percent fixation time spent on the Front view ($M=19.0$, $SD=11.7$) was significantly different than the time spent on the Rear view ($M=15.4$; $SD=9.6$) and between the Side view ($M=18.0$; $SD=10.3$) and the Rear view ($M=15.4$; $SD=9.6$). There was no significant difference between Front ($M=19.0$, $SD=11.7$) and Side view fixation times ($M=18.0$; $SD=10.3$).

Table 3: Average percentage of fixation time for the whole study (all cars and all questions); F-R and R-S are significant.

F	R	S
19.04	15.4	17.97

The data in Figure 10 shows the average percentage of fixation time per car and per question along with the standard error for each.

For most of the cars, participants spent most of the time on the Side view to evaluate 'Sporty-Conservative', 'Basic-Luxurious' and 'Calm-Exciting'. There is no clear indication of the view that participants spent most time on while evaluating 'Exterior Appearance' or 'Appeal', but they spent more time on Front and Side views as compared to the Rear View. There is also no clear indication of a preferred view while evaluating 'Traditional-Innovative'. As already indicated, it is clear that participants spent most time on the Front view while evaluating 'Brand'.

The next section highlights the conclusions and directions for future work.

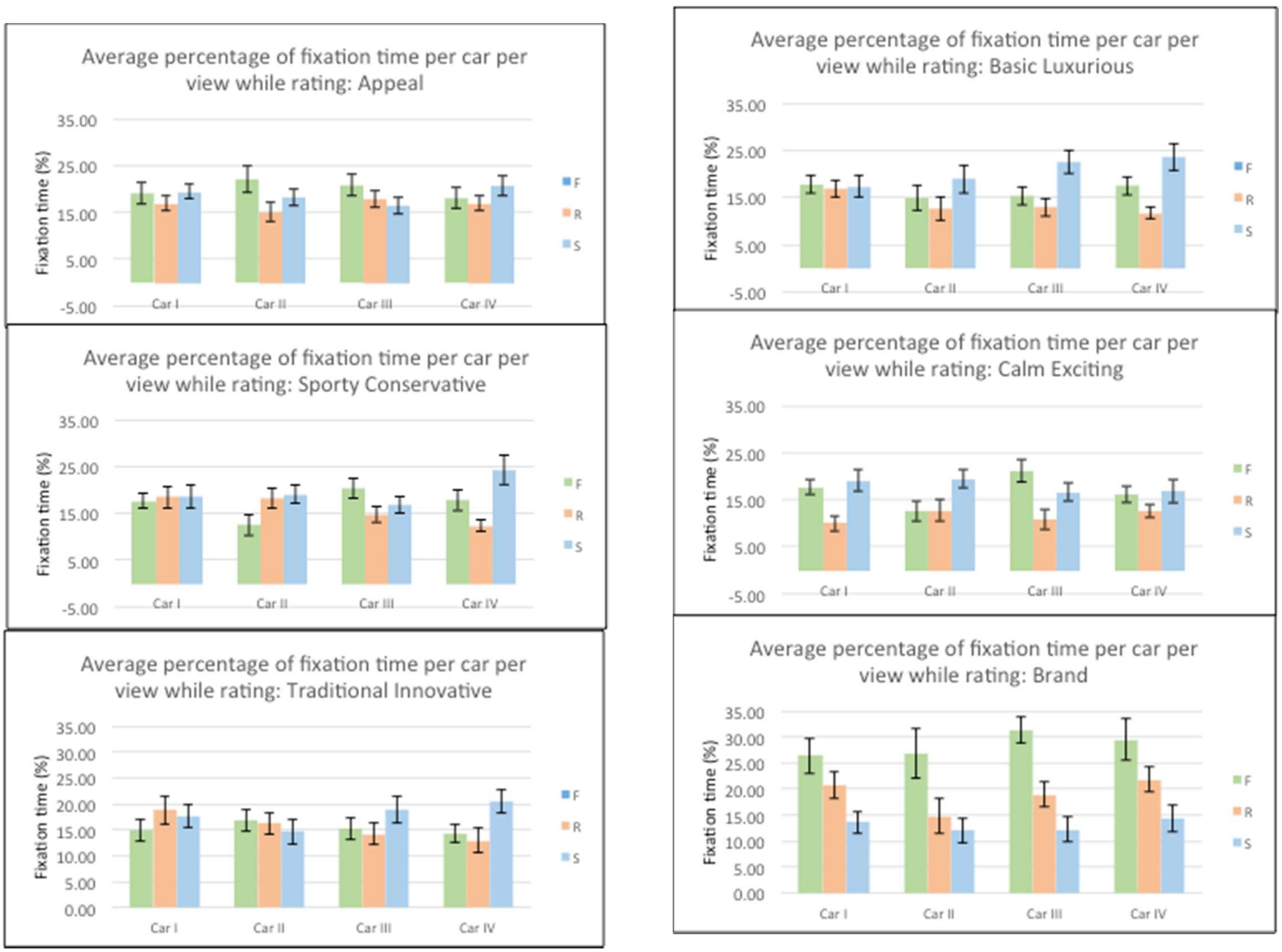


Figure10: Average percentage of fixation time for each car, each question and each view

5 CONCLUSIONS AND FUTURE WORK

Eye tracking data proved to be informative in understanding where people look to find relevant information. When looking for brand cues, respondents tended to focus on the front area. This information is valuable for designers who are seeking insights on what viewing areas may be most relevant for customers. In addition, doing this using 2D images is more cost effective for the designer. When looking for sportiness cues respondents tended to focus on the side, and to a lesser extent the front area. This is valuable to designers who are seeking to express sportiness. Cues for luxuriousness tended to be searched for in the side view. Thus, a designer might experiment with cues shown in this view when seeking to inject luxury into the design. Even non-significant findings are of value to designers. Knowing that patterns of visual inspection do not differ by view when people judge either innovativeness or appeal, a designer may choose not to limit excursions into innovativeness or appeal to just the front, the side, or the rear. However, since this study focused on cars, it has not been validated to generalize the results to other product or artifacts.

This preliminary study served two functions. It answered basic questions about where people look when seeking information of different, design-relevant dimensions, such as innovativeness or brand in 2D representations of cars. But it also sets the stage for complementing visual search data with other, more emotion-laden data, the ultimate motivation for this line of work. Future studies will use eye tracking in conjunction with measures of emotion and engagement (EEG, GSR, facial decoding) and verbal report as respondents inspect both 2D and 3D stimuli to help us understand if there are differences in the perceptual process for the two formats.

AOIs within each view have not been analyzed in this work as this was limited by the size of the views. There were three views on each slide and the intention was to see which view was receiving more attention. The next phase of this study will include AOIs within each view. Future work may present only one image/view at a time thus allowing better resolution and AOI analyses for the questions. This will provide more detailed information about the features or parts of the car exterior that participants look at while making focused evaluations. This has obvious implications for where a designer will focus their efforts to accomplish different design goals.

ACKNOWLEDGMENTS
We would like to thank Wan-Lin Hu, Shruti Rajan Kizhakkel and other members of the REID lab at Purdue University for their help and support in completing these studies. We would also like to thank General Motors for their collaboration in this project.

REFERENCES

[1] Orsborn, S., Cagan, J., and Boatwright, P., 2009, “Quantifying aesthetic form preference in a utility function,” Journal of Mechanical Design, 131(6), pp.061001-61002.

[2] Orbay, G., Fu, L., & Kara, L. B., 2015, “Deciphering the Influence of Product Shape on Consumer Judgments Through Geometric Abstraction,” Journal of Mechanical Design, 137(8), pp. 081103-81104.

[3] Dahan, E., & Srinivasan, V. 2000, "The predictive power of internet-based product concept testing using visual depiction and animation,” Journal of Product Innovation Management, 17(2), pp. 99-109.

[4] Wassenaar, H. J., and Chen, W., 2003, “An approach to decision-based design with discrete choice analysis for demand modeling,” Journal of Mechanical Design, 125(3), pp. 490-497.

[5] Nagamachi, M., 1995, “Kansei engineering: a new ergonomic consumer-oriented technology for product development,” International Journal of Industrial Ergonomics, 15(1), pp. 3–11.

[6] 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), pp. 101010-101011.

[7] Swamy, S., Orsborn, S., Michalek, J., & Cagan, J., 2007, “Measurement of headlight form preference using a choice based conjoint analysis,” 2007 Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol.6, pp. 197-206.

[8] Yumer, M. E., Chaudhuri, S., Hodgins, J. K., & Kara, L. B., 2015, “Semantic shape editing using deformation handles,” ACM Transactions on Graphics (TOG), 34(4), pp. 86.

[9] Sauer, J. and Sonderegger, A., 2009, “The influence of prototype fidelity and aesthetics of design in usability tests: effects on user behaviour, subjective evaluation and emotion,” Applied Ergonomics, 40(4), pp. 670–677.

[10] Bates-Brkljac, N., 2009, “Assessing perceived credibility of traditional and computer generated architectural representations,” Design Studies, 30(4), pp. 415–437.

[11] Artacho-Ramirez, M. A., Diego-Mas, J. A., and Alcaide-Marzal, J., 2008, “Influence of the mode of graphical representation on the perception of product aesthetic and emotional features: An exploratory study,” International Journal of Industrial Ergonomics, 38(11), pp. 942–952.

[12] Nisbett, R. E., & Wilson, T. D., 1977, “Telling more than we can know: verbal reports on mental processes,” Psychological Review, 84(3), pp. 231.

[13] Payne J. W., Bettman, J. R., and Johnson, E. J., 1993. The adaptive decision maker. Cambridge University Press, Cambridge, UK.

[14] Montgomery, H., and Svenson, O., 1989, “A think aloud study of dominance structuring in decision processes,” Process and Structure in Human Decision Making, pp. 135–150.

[15] Billings, R. S. and Marcus, S. A., 1983, “Measures of compensatory and noncompensatory models of decision behavior: Process tracing versus policy capturing,” Organizational Behavior and Human Performance, 31(3), pp. 331–352.

[16] Glockner, A. and Betsch, T., 2008, “Multiple-reason decision making based on automatic processing,” Journal of Experimental Psychology: Learning, memory, and cognition, 34(5), pp: 1055.

[17] Glockner, A. and Herbold, A., 2011, “An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes,” Journal of Behavioral Decision Making, 24(1), pp. 71–98.

[18] Just, M. A. and Carpenter, P. A., 1976, “The role of eye-fixation research in cognitive psychology,” Behavior Research Methods & Instrumentation, 8(2), pp. 139-143.

[19] Goldberg, J. H. and Kotval, X. P., 1999, “Computer interface evaluation using eye movements: methods and constructs,” International Journal of Industrial Ergonomics, 24(6), pp. 631–645.

[20] Loftus, G. R. and Mackworth, N. H., 1978, “Cognitive determinants of fixation location during picture viewing,” Journal of Experimental Psychology: Human Perception and Performance, 4(4), pp. 565–572.

[21] Madsen, A., 2013, “Studies of Visual Attention in Physics Problem Solving,” Ph.D. thesis, Kansas State University, Manhattan, KS.

[22] Djamasbi, S., Siegel, M., Skorinko, J., & Tullis, T., 2011, “Online viewing and aesthetic preferences of generation y and the baby boom generation: Testing user web site experience through eye tracking”, International Journal of Electronic Commerce, 15(4), pp. 121-158.

[23] Meissner, M., & Decker, R., 2010, “Eye-tracking information processing in choice-based conjoint analysis”, International Journal of Market Research, 52(5), pp. 592-611.

[24] Reid, T. N., MacDonald, E. F., & Du, P., 2013, “Impact of Product Design Representation on Customer Judgment,” Journal of Mechanical Design, 135(9), pp. 091008-91013

[25] Du, P., & MacDonald, E. F., 2014, “Eye-Tracking Data Predict Importance of Product Features and Saliency of Size Change,” Journal of Mechanical Design, 136(8), pp. 081005-81006.

[26] Wedel, M., & Pieters, R., 2008, “A review of eye-tracking research in marketing”, Review of marketing research, 4(2008), pp. 123-147.

[27] Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A., 2011, “Search dynamics in consumer choice under time pressure: An eye-tracking study”, The American Economic Review, 101(2), pp. 900-926.

[28] Khushaba, R. N., Wise, C., Kodagoda, S., Louviere, J., Kahn, B. E., & Townsend, C., 2013, “Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking”, Expert Systems with Applications, 40(9), pp. 3803-3812.

[29] Du, P., & MacDonald, E. F., 2015, “Products' Shared Visual Features Do Not Cancel in Consumer Decisions”, Journal of Mechanical Design, 137(7), pp. 071409.

[30] Du, P., & MacDonald, E., 2015, “Eye-tracking Aids in Understanding Consumer Product Evaluations”, The Psychology of Design: Creating Consumer Appeal, Routledge, pp.301.

[31] Gegenfurtner, A., Lehtinen, E., and Slj, R., 2011, “Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains”, Educational Psychology Review, 23(4), pp. 523-552.

[32] Pernice, K. and Nielsen, J., 2009, “Eye tracking Methodology: How to Conduct and Evaluate Usability Studies Using Eye tracking,” Nielsen Norman Group Technical Report.

[33] Coates, D. 2003 “Watches Tell More Than Time: Product Design, Information, and the Quest for Elegance”, New York: McGraw-Hill.

[34] Osgood, C. E., Tannenbaum, P. H., and Suci, G. J. 1957, “The Measurement of Meaning”, Urbana: University of Illinois Press.

ANNEX A

THE DIFFERENT CAR STIMULI USED IN THE STUDY

The various car stimuli used in the study are shown below.

Car I

Car III

Car II

Car IV