Influence of Detailed Photographs of Product on Customer's Purchase Decision

Sanjay Kumar Ranganayakulu
Clemson University
Dept. of Industrial Engg.
147 Freeman Hall, Clemson, SC 29634
rrangan@clemson.edu

Nikhil Prashant Bendre
Clemson University
School of Computing
Clemson, SC 29631
nbendre@clemson.edu

Shaunak Vivek Natu
Clemson University
Dept. of Industrial Engg.
147 Freeman Hall, Clemson, SC 29634
snatu@clemson.edu

Dr.Andrew T. Duchowski
Clemson University
School of Computing
100 McAdams Hall, Clemson, SC 29634
andrewd@cs.clemson.edu

ABSTRACT
The purpose of this experiment was to study the influence of detailed photographs of the product against the review of the product. The product taken into consideration was an apartment and the medium used to view the apartment was a webpage layout. The experiment used a two by two design where one independent variable was level of detail photographs and the other was apartment quality. The qualities of reviews were kept at a constant level. Twenty one participants were made to view the four different layouts of the apartment rental webpage. All the subjects were eye-tracked to determine where and how long they saw the photographs and read the reviews in various layouts. It was found that detailed pictures influence participant's preference of an apartment more than the reviews. Participants tend to spend more time to make a decision if there are less number of pictures than detailed pictures, since they have to rely more on the reviews.

Author Keywords
Eye-tracking, Web usability, E-commerce, Product marketing.

INTRODUCTION
The growing influence of the Internet in our everyday life has reached to such an extent that not only does the web content have to be good but also the usability needs to be excellent. Estimate from 2006 indicate that the average U.S based user viewed 120 web pages per day. [5]

The e commerce industry is flourishing very fast and now-a-days most of the sales take place via the Internet. In such a scenario, it has become important both for the consumer and the seller/marketer to benefit from the eye tracking results which indicate which part of the webpage has the most visual attention.

When people buy products off the Internet most of the times they tend to read the reviews about the product. We conducted a study to find how detailed pictures of a product can influence a consumer's decision. The study consists of presenting two templates to the users. One template consisted of just the user reviews and other consisted of the user reviews and detailed pictures of the product both. The product we are taking into consideration is an apartment listing. We are proposing a within group experiment.

BACKGROUNG
The concept of observing eye-movements for implying what humans think about a system has been introduced a century before computers were widely used [8]. During the initial stages of developing equipment's to track peoples eye movements they were all invasive like making people wear eye lenses sticking out electrode from them. In 1901, Dodge and Cline [1] developed the first precise, non-invasive eye tracking technique using light reflected from the cornea of the eye. Pioneers in the field of HCI and usability like Fitts [3] used these methodologies to device theories like Fitts law.

Since 1970's there has been great advances in both eye tracking technology and psychological theory to link eye tracking data to cognitive processes[8]. There has been a tremendous development in the eye-tracking technology and devices since then. In recent years due to proliferation of personal computers and extensive day to day use of internet and websites, usage of eye trackers to conduct usability research has increased. Many researchers like Neilsen [10] and Cutrell [5] have conducted extensive

CHI 2012, May 5-10, 2012, Austin, TX, USA.

general web designing layout usability studies and came up with general layout guidelines for designing usable websites. However, there are very few examples of research regarding people's eye movements as they integrated both textual and visual elements in decision making context. As noted by Rayner et al. [11], this may be due to the fact that the bulk of this research has probably been conducted by advertising agencies and the majority of it is not made readily available to the scientific community [8].

Few studies closely related to our topic are discussed. Faraday and Sutcliffe [2] conducted an experiment to investigate how the correlations between textual and pictorial elements are important to make it easy for the users to understand the content. They recommended that "co-reference between an image and a text" [2, pg. 29] should be carefully constructed to ensure that the maximum amount of encoded information is passed along to the viewer. Their findings revealed that, participants sought to link textual descriptions to visual depictions in a simple manner and if the link wasn't clear, participants often became confused as to how the two channels could be synthesized into a coherent whole. This work was based upon earlier research conducted by Hegarty [6,7]. As her participants viewed instructional diagrams comprised of both visual and textual directions, they constructed representations of the instructional material that were mainly text based. Across analyses, imagery was found to supplement an initial and lasting impression that was decoded from the textual material. Thus these above researches prove that detailed images reinforcing the text help users to understand the context better.

The remaining research was generated by investigations of people's interactions with advertising materials. This work is mainly interested in what pieces of an advertisement are most likely to capture a user's interest. Work by Fox et. al. [4] and Lohse [9] are representative examples. A final study of particular interest was conducted by Rayner et al.[11]. That work used magazine advertisements as source material for participants conducting searching related tasks. Participants were directed to, for example, "decide which skin lotion and body wash to buy", from a series of advertisements that contained both pictorial and textual elements. Results indicated that participants spent a majority of their time looking at the textual portions of the advertisement. While this was true, it was hypothesized that participants spent more time on the text due to the fact that images are more quickly and easily decoded than words. Their work also pointed to the fact that eye movement patterns are heavily affected by task. The current study is intended to address the research question how detailed photographs of a product might influence user's decision making in buying the product and whether users care to read the product description and reviews if they could see detailed photographs of the product.

HYPOTHESIS

It is hypothesized that if a set of detailed photographs of an apartment are presented then less people would bother to read the reviews.

It is also hypothesized that people would respond positively after viewing a page with detailed set of photographs of the product versus the webpage with only one photograph of the product.

EXPERIMENT METHOD

Participants

Twenty-one students from Clemson University were recruited through an email and/or verbal invitation describing the study. Students expressing an interest in participating were pre-screened via questionnaire to determine their eligibility: participants should have prior experience using the Internet for a minimum of one year and normal (or corrected to normal) visual acuity, and no history of migraine headaches or epilepsy. Those with bifocals and contact lenses were not further considered, due to limitations of the eye tracker. This pool of 21 participants was shown all four layouts of the website in random order. Two layouts showed detailed photographs of bad and good apartment and the other two showed one photographs of bad and good apartment.

Experimental Design

The experiment used a two level within-subject repeated measures design as shown in Table 1. Each of the two conditions, or levels, of the independent variables, level of detail and apartment quality of webpage layout, will use the same webpage structure and quality of reviews except for the photographs section. Participants will be assigned to conditions in a randomized way using tobii studio to minimize order effects. The data will be collected from each participants view on AOI of the layouts that is picture region AOI and reviews region AOI, and subsequently statistically analyzed.

The dependent variables in this experiment include objective and subjective measures of performance for each participant. The experimental study will be conducted in single sessions and objective measures measured in this session are the number of fixations (fixation counts) in each AOI made by the participants, and the dwell time (fixation duration) spent by participants in each AOI.

Table 1. Within subject design with 2 x 2 levels

	Level of detail (independent variable 1)	
	High	Low
Apartment Quality (independent variable 2)	Webpage with detailed photographs of the apartment	Webpage with no detailed photographs of the apartment
Bad		
Good		

Subjective data was obtained through a preference ranking post experiment questionnaire administered to the participants at the end of the experimental study. The data collected for each dependent measure was statistically analyzed using a two-way analysis of variance. The locus of significant differences, if any, was analyzed using Tukey's honestly significantly different (HSD) post-hoc test.

Testing Environment

The study will be conducted in the Virtual Reality and Eye-tracking Laboratory at Clemson University. The experimental set-up will consist of an eye-tracker enabled computer, table, chair, paper and pencil. The eye-tracker used to conduct this study will be the Tobii ET-1750 (Tobii Technology, Inc., Falls Church, VA). The eye-tracker is embedded in a TFT 17" monitor that will run at a resolution of 1280x1024. The system samples eye position data at 50Hz with an accuracy of about 0.5 degrees. The software that will be used to conduct the experiment will be Tobii's Studio (v2.2.8). The computer screen will display a web-page of an apartment finder website.

RESULTS

We used SPSS 19 for data analysis. For both of our objective dependent variable fixation duration and fixation counts, we used two-way ANOVA with a 95% confidence interval to determine significance. We had two independent variable (IV) level of detail and apartment quality. Each of these IV's had two levels, high level of detail represented more photos of the apartment and one overall photo for low level of detail. Similarly good apartment quality had photos of a good apartment and bad apartment quality had photos of a bad apartment. However the reviews of the apartment were written in the same style and depicted the same inference on the readers. The area of interest (AOI) were split up as pictures section versus review section in all the stimuli's as shown in the Figure 1.

Figure 1: Representation of Stimuli AOI's

So, in order to analyze these data collected in various stimuli AOI's we used two-way ANOVA. The interaction effects among the IV's and the AOI's with respect to IV's were analyzed by creating multiple spilt data sheets and performing two-way ANOVA individually on them. Since we split data and analyzed we didn't need separate post-hoc test to identify the locus of significant difference among IV's. The multiple split data sheets were created by holding one IV as constant and varying the others, for e.g. we kept level of detail (high vs low) constant and changed the other IV's data accordingly in this format as shown in Figure 2.

Participant	Apartment_detail	AOI	Fixation_duration_high	Fixation_counts_high	Fixation_duration_low	Fixation_counts_low
1.00	good	picture	4.73	15.00	.64	2.00
2.00	good	picture	3.39	9.00	4.14	13.00
3.00	good	picture	3.23	11.00	3.13	9.42
4.00	good	picture	3.63	12.00	6.40	20.00
5.00	good	picture	8.16	31.00	2.08	7.00
6.00	good	picture	4.98	20.00	4.05	14.00
7.00	good	picture	13.64	36.00	2.31	7.00
8.00	good	picture	5.33	21.00	1.22	4.00
9.00	good	picture	5.50	19.00	.98	4.00
10.00	good	picture	40.13	137.00	8.43	28.00
11.00	good	picture	2.02	6.00	1.38	3.00
12.00	good	picture	7.48	30.00	2.19	6.00
13.00	good	picture	.62	3.00	2.91	10.00
14.00	good	picture	41.37	120.00	5.40	16.00
15.00	good	picture	7.08	21.00	5.48	14.00
16.00	good	picture	3.21	7.00	4.35	4.00
17.00	good	picture	3.62	13.00	.54	2.00
18.00	good	picture	8.93	34.00	2.77	8.00
19.00	good	picture	6.93	22.00	3.13	9.42
20.00	good	picture	7.38	24.00	2.61	10.00
21.00	good	picture	4.07	14.00	1.62	5.00
1.00	good	review	6.97	23.00	5.61	16.00
2.00	good	review	3.68	12.00	7.91	25.00
3.00	good	review	.88	3.00	10.01	32.00
4.00	good	review	1.66	8.00	10.44	35.00
5.00	good	review	17.96	72.00	12.50	43.00
6.00	good	review	26.16	69.00	8.65	28.00
7.00	good	review	.62	2.00	7.23	20.00
8.00	good	review	.72	3.00	.66	3.00
9.00	good	review	24.54	50.00	4.00	14.00
10.00	good	review	60.52	167.00	45.68	162.00
11.00	good	review	4.98	19.00	13.77	45.00
12.00	good	review	5.18	19.00	13.80	49.00
13.00	good	review	28.74	87.00	2.89	10.00

Figure 2: High vs low level of detail split data sheet

Level of detail (High vs Low)

After performing the above describe analysis on the constant high vs low, we found out that the fixation duration on the low level of detail's picture AOI was significantly ($p<0.001$) less than the review AOI in both good and bad apartment quality layout. There was no significant difference among the bad quality pictures and good quality pictures in the picture AOI, similarly also for the review AOI. This is provided in the Figure 3 interaction plot with error bars. Even fixation count on the low level of detail layouts followed the same trend as the fixation duration.

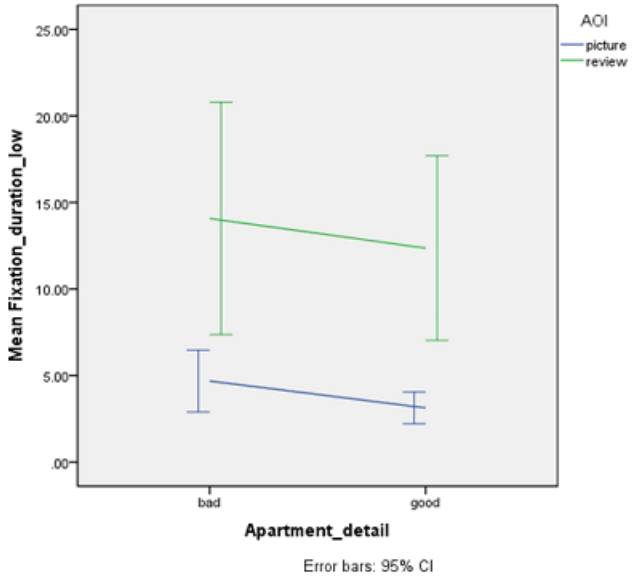

Figure 3: Interaction plot for low level of detail layouts fixation durations.

There was no significant difference found among fixation durations of high level of detail AOI's with respect to good and bad quality of apartment.

Quality of Apartment (Good vs Bad)

Two-way ANOVA was performed on the fixation duration and fixation counts data stored in the good vs bad apartment quality file. When the apartment quality is good fixation duration is significantly less ($p<0.001$) on photos AOI than reviews AOI. Even for the bad apartment quality it's the same, people spend more time in reading reviews if there are less number of photos. This interaction graph is provided in figure 4 and 5.

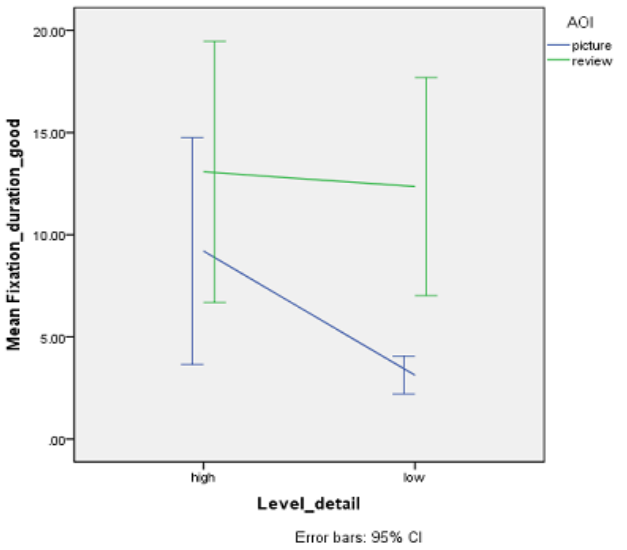

Figure 4: Interaction plot for good apartment quality fixation durations.

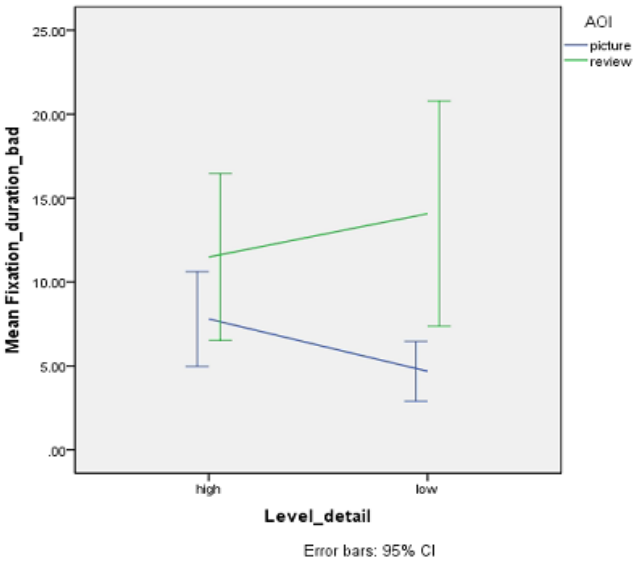

Figure 5: Interaction plot for bad apartment quality fixation durations.

The fixation counts data also showed a similar trend, there were significantly more fixations counts on the review AOI if there are less detailed photographs of the bad or good quality apartment.

Area of Interest (Picture vs Review)

Similar statistical analysis was performed by taking AOI's as constant axis and varying other IV's. Analysis on the picture AOI showed that fixation duration is significantly less ($p<0.001$) when less number of pictures are shown to the participants. This is shown in the Figure 6. For the reviews AOI fixation duration is not significant in any of the IV's since reviews were maintained at a constant level and people read through them. This is shown in the Figure 7.

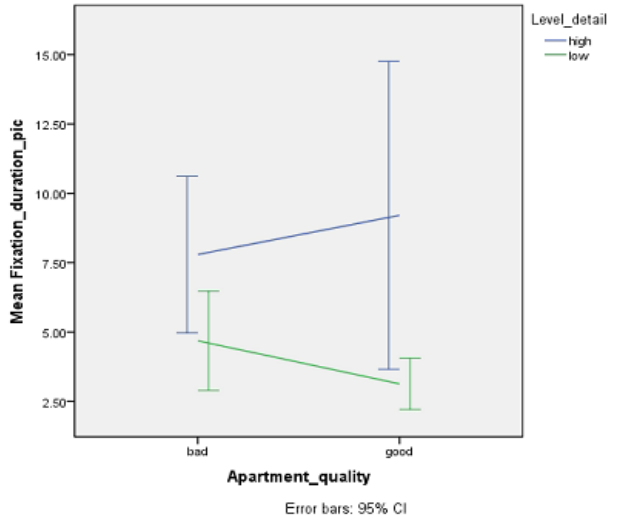

Figure 6: Interaction plot for Picture AOI fixation durations.

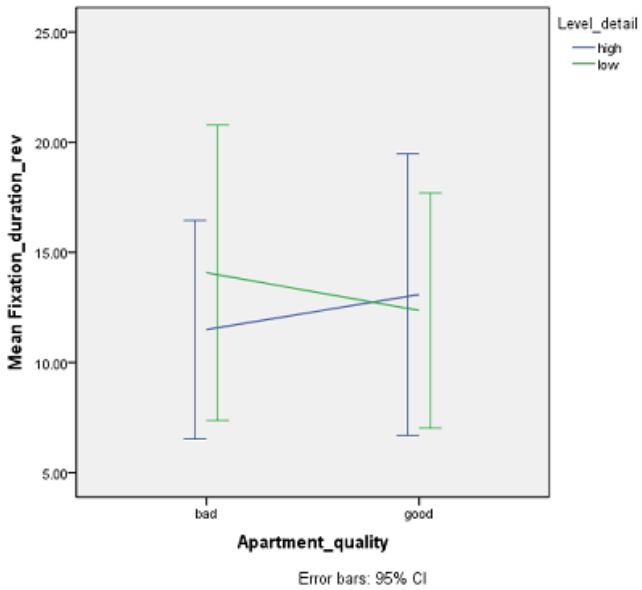

Figure 7: Interaction plot for Review AOI's fixation durations.

The fixation count also followed a similar trend as the fixation duration values. That is more pictures having more fixation counts than less picture layouts. Review AOI's having the same level of fixation counts throughout all conditions.

Preference Ranking

The subjective dependent measure that was collected was the ordered preference ranking of the apartment shown to the participants. The distribution of the highest preferred layout and the overall rank distribution among layouts are shown in the Figure 8 and 9 respectively.

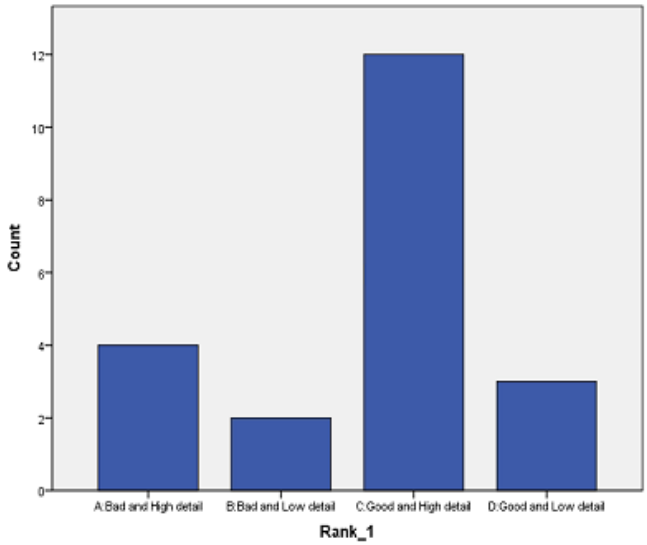

Figure 8: Distribution of number of participants ranking each layout highest in terms of preference.

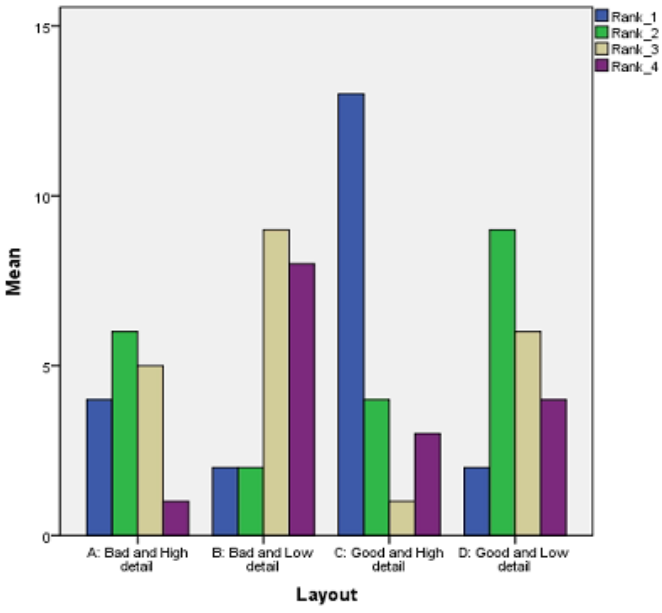

Figure 9: Overall preference ranking distribution.

Since the ranking data was not normally distributed we used Friedman non-parametric test to analyze the data. We then used pairwise Wilcoxon signed-rank tests with Bonferroni correction (significant if $p < 0.025$) to see which layouts differed significantly.

The results of the Friedman test showed that layout C was ranked the highest among user preference with mean rank of 1.76 (lower the mean rank the better) followed by layout D (2.48), A (2.71) and B (3.05). Using the above mention post-hoc test we found that Layout C is significantly different than B ($Z = -2.398$, $p = 0.016 < 0.025$). All the other layouts were not significantly different from each other.

The fixation duration on each of the stimuli can be visually represented using heat maps. The heat maps comprise of the comprehensive data for all twenty one participant's fixation durations on each of the web layouts as shown in Figure 10,11,12,13. We could observe in these heat maps that when there are detailed photos in the layout, participants tend to look at them more, represented by red spots in them. In the opposite sense if there were fewer pictures participants tend to spend less time on pictures and read reviews for longer durations.

Figure 10: High detail, bad quality apartment layout (A)

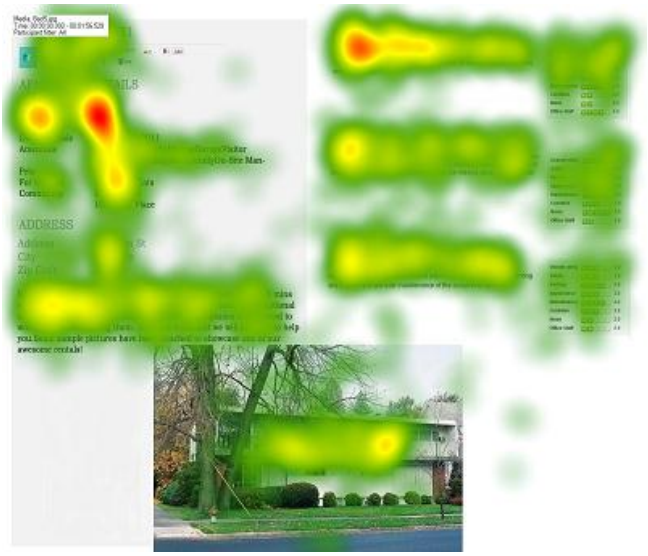

Figure 11: Low detail, bad quality apartment layout (B)

Figure 12: High detail, good quality apartment layout (C)

Figure 13: Low detail, good quality apartment layout (D)

DISCUSSION

The objective and subjective measures were statistically analyzed as explained in the above section. To be consistent with the result section we discussed them in three separate sections as shown below.

We observed that in case of high level (more pictures) against low level (less pictures) of detail, the results for the study group indicate that for a high level of detailed pictures, mean fixation duration and count was insignificant even when we have a set of high and bad quality of pictures while keeping the quality of reviews constant. So, the participants still tend to see the high level of detailed pictures, even if the quality of the apartment is good or bad. Similarly, for low level of detailed pictures, the participants tend to spend more time reading the reviews when it comes to a bad quality. One possible explanation for this maybe that they try to make decisions based on reviews than the low level of detailed pictures presented to them. When it comes to good quality pictures for low level of detailed pictures they try to make decision based on picture rather than the reviews.

When we compare it based on good or bad apartment quality layouts, if there were less number of pictures, mean fixation duration and count was significantly less on pictures than on reviews. This might be because the participants spent more time reading the reviews since there were less number of pictures to make a decision. When the numbers of pictures were more, they spent less time in viewing them than reading reviews but that was also not significantly different as people might process pictures faster than reading the text reviews [12]. So, in general when participants are given more detailed pictures they might read less reviews and make a decision based solely on the given detailed set of pictures.

Based on the areas of interest (AOI), picture AOI indicates that mean fixation duration and count is significantly less when less number of pictures are shown to the participants and for the reviews AOI mean fixation duration and count is not significant since reviews were maintained at a constant level. We may infer from this that if the participants are given less number of pictures they are more likely to depend on reviews for decision making process but at the same take a longer time in reading those reviews and probably will be unsure of their decision since it only depends on the review and not the pictures. The participants may try to correlate the less number of pictures with the reviews and so the scan paths indicate their eye movement from the reviews to the fewer amounts of pictures and then back to reviews. This may show us that they are trying to feel confident about their decision based only on the review. When the picture AOI contains a set of detailed pictures, the mean fixation time is less than the review AOI and so it could mean that the participants were sure of their decisions.

CONCLUSION

This study explored the role of pictures and text reviews in influencing participant's preference based on viewing a webpage of apartment details. The results of the study which consisted of a small number of participants indicated that when provided with less detailed pictures of the apartment, the participants took more time in reading the reviews in order to come to a decision about the preference of the apartment. This was in accordance with our first hypothesis. The preference ranking data inferred that the detailed pictures of an apartment could influence people to prefer an apartment even though both have the same quality of reviews.

FUTURE WORK

In this study, we had kept the reviews constant for both of the independent variables. We could change the reviews also and perform the experiment for a set of good and bad reviews. We have considered only an apartment layout as a stimulus but we could extend it for different products as well. By considering a wider range of products we could get a more conclusive result and thus make a better classification between pictures and review when it comes to online product marketing.

ACKNOWLEDGMENTS

The authors wish to thank sincerely Mr.Drew Parker and Mr.Barrett Wood of Vidaloo,LLC for letting us use their website www.rentalchatter.com as a model for creating stimuli for the experiments. We would also like to place on record our profound gratitude and sincere thanks for the excellent guidance we received under Dr.Andrew Duchowski for this project.

REFERENCES

1. Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*, 145–157.
2. Faraday, P. & Sutcliffe, A. (1998) Making Contact Points between Text and Images. *Proceedings of the sixth ACM International Conference on Multimedia*. Bristol, UK. pp. 29-37.
3. Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.
4. Fox, R., Krugman, D., Fletcher, J. & Fischer, P. (1998) Adolescents' Attention to Beer and Cigarette Print Ads and Associated Product Warnings. *Journal of Advertising*, 27. pp. 57-68.
5. Georg Buscher, Edward Cutrell, and Meredith Ringel Morris. 2009. What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In *Proceedings of the 27th international conference on Human factors in computing systems* (CHI '09). ACM, New York, NY, USA, 21-30. DOI=10.1145/1518701.1518705 http://doi.acm.org/10.1145/1518701.1518705
6. Hegarty, M. (1992) The Mechanics of Comprehension and Comprehension of Mechanics. In *Eye Movements and Visual Cognition: Scene Perception and Reading*. ed. Rayner, K. New York: Springer-Verlag. pp. 428-448.
7. Hegarty, M. (1992) Mental Animation: Inferring Motion from Static displays of Mechanical Systems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18. pp. 1084-1102.
8. Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section commentary). In J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. Oxford: Elsevier Science.
9. Lohse, G. (1997) Consumer Eye Movement Patterns on Yellow Page Advertising. *Journal of Advertising*, 26. pp. 61-73.
10. Nielsen, J. (2006). F-Shaped pattern for reading Web content.Jacob Nielsen's Alertbox for April 17. Retrieved April 17, 2006 from http://www.useit.com/alertbox/reading_pattern.html
11. Rayner, K., Rotello, C., Stewart, A., Keir, J., & Duffy, S. (2001) Integrating text and pictorial information: eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, 7(3). pp. 219-226.
12. Anthony Hughes, Todd Wilkens, Barbara M. Wildemuth, and Gary Marchionini. 2003. Text or pictures? an eyetracking study of how people view digital video surrogates. In Proceedings of the 2nd international conference on Image and video retrieval (CIVR'03), Erwin M. Bakker, Michael S. Lew, Thomas S. Huang, Nicu Sebe, and Xiang Zhou (Eds.). Springer-Verlag, Berlin, Heidelberg, 271-280.