On Saliency, Affect and Focused Attention

Lori McCay-Peet
Dalhousie University
Halifax, Nova Scotia, Canada
mccay@dal.ca

Mounia Lalmas
Yahoo! Research
Barcelona, Spain
mounia@acm.org

Vidhya Navalpakkam
Yahoo! Research
Santa Clara, California, USA
nvidhya@yahoo-inc.com

ABSTRACT
We study how the visual catchiness (saliency) of relevant information impacts user engagement metrics such as focused attention and emotion (affect). Participants completed tasks in one of two conditions, where the task-relevant information either appeared salient or non-salient. Our analysis provides insights into relationships between saliency, focused attention, and affect. Participants reported more distraction in the non-salient condition, and non-salient information was slower to find than salient. Lack-of-saliency led to a negative impact on affect, while saliency maintained positive affect, suggesting its helpfulness. Participants reported that it was easier to focus in the salient condition, although there was no significant improvement in the focused attention scale rating. Finally, this study suggests user interest in the topic is a good predictor of focused attention, which in turn is a good predictor of positive affect. These results suggest that enhancing saliency of user-interested topics seems a good strategy for boosting user engagement.

Author Keywords
User engagement, saliency, focused attention, positive affect, user interests, news entertainment

ACM Classification Keywords
H5.m. Information interfaces and presentation: Miscellaneous.

General Terms
Human Factors, Experimentation, Measurement

INTRODUCTION
In web applications, *user engagement* refers to the quality of the user experience emphasizing the positive aspects of the interaction, in particular the phenomena associated with being captivated by the application, and wanting to use it frequently. User engagement is a key concept in the design of web applications, motivated by the observation that successful applications are not just used, they are engaged with. In a world full of choices, it is essential that technology providers do not just design systems but that they design engaging experiences [21].

User engagement can be improved by serving not only the most relevant content/functionality to users, but also by displaying it in a *visually catchy way*. The sensory and visual appeal of a web page is an important factor for engagement [18]. In the context of online shopping, web searching, educational webcasting and video games [18], web page style has been related to factors such as screen layout, graphics and the use of design principles such as symmetry and balance. In the context of multimedia design [8], web page style (here with respect to media quality) has been related to *positive affect* and to promote *focused attention*.

Positive affect and focused attention are important characteristics of user engagement [1]. Focused attention refers to the exclusion of other things [18]; while affect relates to the emotions experienced during the interaction. Jennings et al. [8] argue that affective experiences are intrinsically motivating and that, in relation to engagement, an initial affective hook can induce a desire for exploration or active discovery. One important, albeit less studied, aspect, of web page style, is the visual catchiness of items, known as *visual saliency.* The aim of this work is to study the effect of saliency on focused attention and positive affect.

Research in cognitive psychology and neuroscience show that when users see a display or visually search for a target object in a cluttered display, their attention and eye gaze are drawn to salient objects that appear visually different from the rest of the display [24]. For example, big, bolded, italicized and colored text surrounded by plain text appears visually catchy or salient and draws attention, due to the difference in size, orientation and color. More generally, differences in features such as color, brightness, edge orientation, size, motion and flicker are known to draw attention, independently of the user goal or intent [26].

While saliency is known to attract eye gaze in displays, we do not know if it leads to a better user affect and focused attention. To answer this question, we carried out a study within the news entertainment domain, where we measured, both quantitatively and qualitatively, affect and focused attention for a large sample of participants drawn from Amazon Mechanical Turk [13].

We carefully varied the saliency of the headlines in a news page from non-salient to salient, while keeping everything else fixed, and tested whether increasing saliency can help improve user task performance, and whether it impacts focused attention and affect. To help calibrate saliency of the headlines, we turn to a model of saliency [7] that takes in a screenshot of the news page as an input and predicts the salient hotspots that will likely draw user attention in the first few seconds of visiting that page.

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

The paper is organized as follows. We start with a description of our methodology. Next, we present our results, which we then discuss and relate to results from the literature. We end with implications and suggestions for future work.

METHODOLOGY

Our goal was to explore the impact of saliency on positive affect and focused attention. We selected the entertainment news domain knowing that the internet is a common way for people to get their entertainment news [20]. To ensure control of the saliency of web pages we were restricted in terms of the tasks to create. However, these restrictions allowed a more focused study at a more granular level, albeit at the expense of a less natural reading experience.

Participants drawn from Amazon Mechanical Turk were asked to complete an online study. There were two conditions for this study, salient and non-salient.

- *Salient:* this group is shown only entertainment web page images in which the task-based information is salient
- *Non-salient:* this group is shown only entertainment web page images in which the task-based information is not salient

The following hypotheses were tested using both quantitative and qualitative measures:

- *H1: Saliency of task-relevant information leads to greater task performance, both actual and perceived*
- *H2: Affect improves when the task-relevant information is salient compared to when it is not salient.*
- *H3: Saliency of task-relevant information leads to higher levels of focused attention.*

We were also interested in the interaction between participants' level of interest in the entertainment web page content with affect and focused attention, but due to its exploratory nature, we did not formulate hypotheses.

Saliency model

We used the saliency model of visual attention developed by Itti and Koch [7], inspired by how the primate visual cortex processes information to drive attention while freely viewing the display, in the absence of user goal. Extensions of this model have shown to predict user attention during visual search for a target object in a display [14], while watching video [4], and playing exploratory video games [22]. While the model is precognitive and agnostic to the semantics and user biases, it is a good predictor of attention in the first few seconds of seeing the display and hence serves as a good starting point to predict what might catch the users' eye as soon as they see the display.

In the first stage, the model extracts a set of biologically inspired features, such as color, brightness intensity, and edge orientation at various spatial scales. Feature difference maps are computed by determining the pair-wise difference in feature response at every image location between spatial scales. This results in 42 feature difference maps. In the second stage, the model mimics competitive interactions between neutrons by applying a difference of Gaussian filter (mimics the narrow self-excitatory connections between nearby neutrons and the broad inhibitory connections between far away neurons). This is followed by rectification to remove negative values.

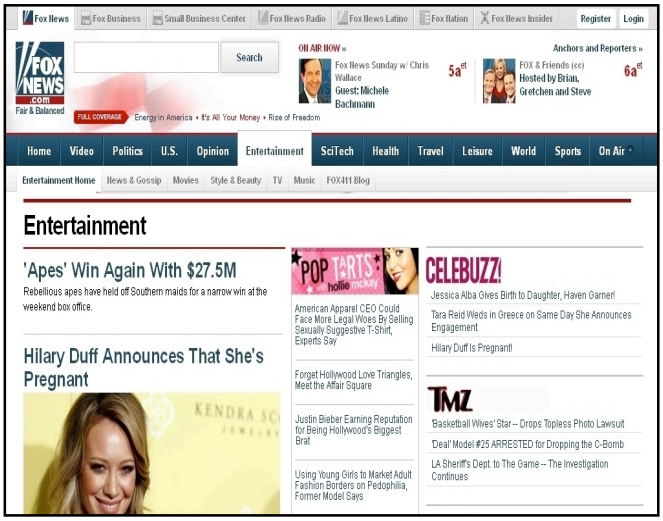
A few iterations of such nonlinear competitive interactions results in highlighting only the salient hotspots in the display in each feature dimension i.e., those display regions whose features are different from the rest of the display. Finally, in the third stage, the model linearly combines the saliency maps from the various feature dimensions to a single master saliency map. Attention is then guided in decreasing order of saliency in this map, starting with the most salient hotspot, followed by the next most salient hotspot and so on until the entire display is visited.

Tasks

There were eight tasks in total for each participant to perform. Each task involved finding the latest news or headline on a different celebrity or entertainment topic (e.g., "Find the headline on Brad Pitt on the next page"). Task topics included Justin Bieber (singer), Sam Worthington (actor), Brad Pitt (actor), Halle Berry (actress), Disney (media conglomerate), Britney Spears (singer), David Letterman (television host), Brendan Gleesan (actor). Entertainment news web pages were selected from a variety of sources to ensure diversity of topic and presentation format (e.g., OMG Yahoo!, CinemaBlend.com, BBC Entertainments & Arts). Only the screenshots of the web pages were manipulated by condition. For consistency, the following manipulation criteria were applied:

1. All web page screenshots have a width:length ratio of 2:3
2. Location of headline on the pages is the same between conditions, although it may vary within a condition
3. Though pages may be scrolled, all headlines appear on the visible page (screen resolution of 1024 x768)
4. According to the saliency model of visual attention, headlines in the non-salient condition do not register as salient; those in the salient condition register as salient

Manipulations included increases in font size, changes in color of font or background, and addition or removal of image or icon. For example, Figure 1 shows the Justin Bieber task web page screenshot images. The headline (lower center of each image) has been manipulated in the salient condition; the font size has been increased, the subtitle font color has been altered and two large exclamation points added. To ensure these web page images contained a non-salient and a salient Bieber headline for their respective conditions, these visible page screenshots were uploaded to a program which automatically applies the saliency model of visual attention described in the previous section to the images. The



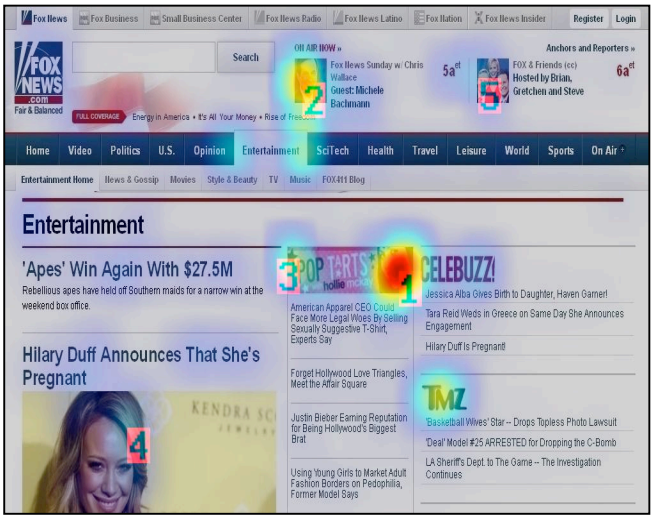
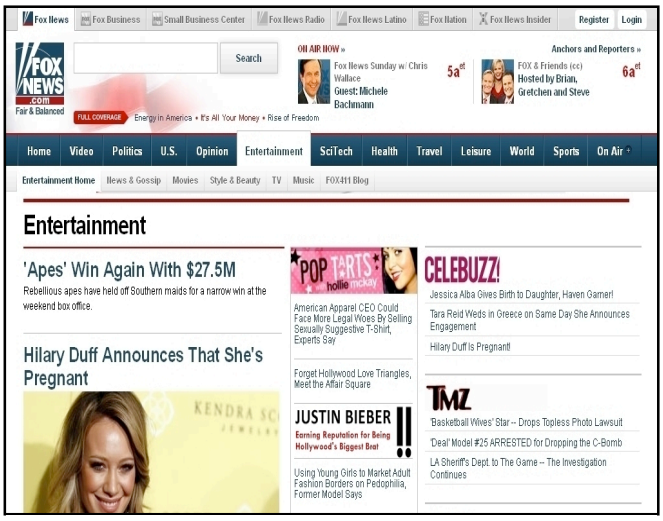


Figure 1. Web page screenshots for the Bieber task (non-salient condition at top, salient at bottom).



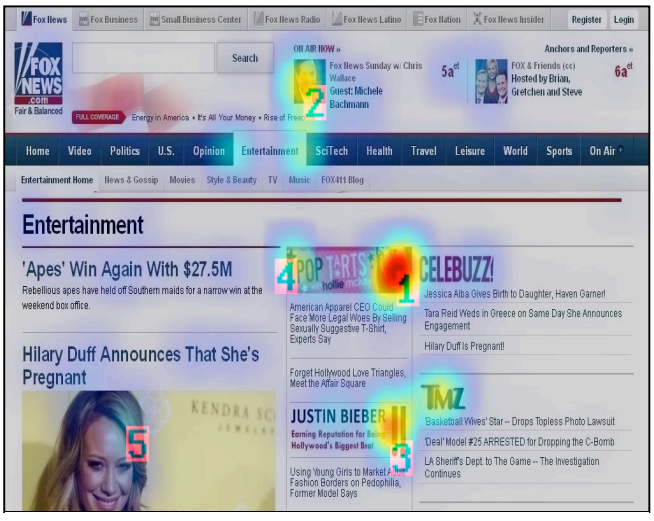


Figure 2. Saliency maps for the Bieber task (non-salient condition at top, salient at bottom)

program outputs a saliency map like those displayed in Figure 2 that identify and number the five most salient spots on the web page images. While the Bieber headline does not register as salient in the non-salient condition, the headline registers as the third most salient item on the page in the salient condition. A score of "1" indicates the feature has the highest level of saliency on the page followed by "2" and so on.

Qualitative and quantitative measures

Affect

Affect was measured pre- and post- task using the Positive and Negative Affect Schedule (PANAS) [25], a 20-item scale that is one of the most commonly used measures of affect. It includes 10 items measuring positive affect (PAS), such as determined, attentive, and enthusiastic and 10 items measuring negative (NAS) including jittery, hostile, and afraid. Participants responded with their agreement on a 5-point scale (very slightly or not at all; a little; moderately; quite a bit; extremely) to the following statement relating to each item: "You feel this way right now, that is, at the present moment."

Affect was also measured by asking users to respond to the question: "Overall, did you feel positive or negative while completing the news headline tasks? [positive, negative, neither positive nor negative]". A follow-up question was designed to understand the reasons behind these feelings (e.g., "Why did you feel positive while completing the news headline tasks?").

Focused attention

Focused attention was measured in a number of ways. First, the 7-item *focused attention* subscale, part of the 31-item user engagement scale [16], was adapted to fit in the context of finding news headlines. It consisted of the following items:

1. I lost myself in this news tasks experience

2. I was so involved in my news tasks that I lost track of time
3. I blocked things out around me when I was completing the news tasks
4. When I was performing these news tasks, I lost track of the world around me
5. The time I spent performing these news tasks just slipped away
6. I was absorbed in my news tasks
7. During the news tasks experience I let myself go

Participants were instructed to respond to each item on a 5-point scale (strong disagree to strong agree): "Based on this experience finding celebrity and entertainment news headlines, please indicate whether you agree or disagree with each statement".

Two questions were included to measure participants' perceptions relating to focus and distraction: "Was it easy to focus on finding the news headlines you were asked to find? [yes, no, somewhat]" and "Did you find anything distracting when you were searching for the headlines? [On the web pages or in your own environment.] [yes, no, somewhat]". These questions were followed with "why" questions to understand reasons for their perceptions.

As perceived time is an integral component of focused attention [16], the actual time spent on the entertainment web page displays looking for headlines was measured as well as perceived time. Perceived time was measured using the question: "Please estimate in minutes the total time you think you spent finding the eight news headlines on the entertainment web pages. [not including time spent reading the questions or selecting multiple choice answers]."

Interest

The participants' level of interest in topics (pre-task) was measured using a 5-point level of interest scale with an option to indicate that they did not know who/what this was. Post-task, interest was measured on a 5-point agreement scale using the following questions: "Though it wasn't possible, I wanted to click on links within the web pages", "I was interested in the content of the web pages" and "I wanted to find out more about the topics that I encountered on the web pages."

Task ease

Perception of task ease was measured based on responses on a 5-point agreement scale to the following two questions: "I am confident I found the headlines I was asked to find" and "It was easy to locate the headlines I was asked to find." Actual time spent finding the headlines was also used as a measure of task ease; the easier the task is to complete, the less time it should take.

Mechanical Turk

The use of Mechanical Turk (MTurk) in research has been explored in detail elsewhere, e.g., [10, 13]. MTurk has been used successfully in the recruitment of participants for external surveys that require longer than your average Mechanical Turk HIT [12], providing a convenient participant pool. To our knowledge, this is the first time that MTurk (or similar platform) was used to study affect and focused attention.

Pilot Studies

We conducted a series of three pilot studies. These pilots enabled us to test the validity of our use of the focused attention subscale, previously applied to the online shopping domain [18]. In the initial pilot study (N=191) confirmatory factor analysis was performed using the Principal component analysis method of extraction and Varimax rotation; a single factor was extracted and Cronbach's alpha (.93) indicated high reliability. The second and third pilot studies further confirmed its reliability; Cronbach's alphas of .90 (N=81) and .88 (N=46) were found. The pilots also allowed us to test the study design and led to a number of significant changes in task, question order and delivery, the types of questions asked, and payment structure.

It was imperative that we control users' view of the web page images as much as possible given that they would be in their own environments rather than in a controlled lab setting. Complaints regarding slow page loading in the first pilot, for example, led us to insert coding to help ensure that web pages pre-loaded. This reduced the chance that participants would see the page slowly appear from top to bottom, thus impacting saliency predictions.

Comments in the original pilot study suggested that some tasks meant to be in the salient condition were considered difficult to spot while some in the non-salient condition stood out. Most of the complaints regarding the non-salient headlines in the salient condition were related to highlighting that impaired readability; new tasks were created with this in mind. The second and third pilots became pared-down versions of the study (e.g., PANAS questionnaire removed) in order to get the tasks "right".

The original pilot study included follow-up questions designed to engage the participants in the content. Comments suggested this was too artificial given the lack of ability to engage with the web pages via links; as well, those with no interest in entertainment news complained vehemently. Given the sensitivity of the PANAS, follow-up questions were removed and replaced with a question that simply asked them to select the correct headline from a list of headlines. The number of tasks was also reduced to eight from ten to reduce emotions not associated with saliency itself. The polarized reactions to entertainment and celebrity news led us to add a few questions relating to interest in the web page content.

Overview of study protocol

The study took place over a period of approximately 12 hours in August, 2011. **Error! Reference source not**

found. outlines the flow of the study from its introduction to Amazon Mechanical Turk Workers and acceptance of the HIT assignment to their return to the Amazon site to submit the HIT after completing the study. Participants were instructed that the study would take approximately 15 minutes to complete, though they would be given 60 minutes between the time they accepted and submitted the HIT. They were informed that they could only participate in this study once and they would not be paid if they had participated in one of the two previous pilot studies that employed the same tasks. Payment for study completion was $1.50. Upon accepting the HIT, participants were randomly assigned to one of 8 surveys by clicking the link to the external survey.

Items in the PANAS and focused attention scale were randomized as were questions in the pre-task interest questionnaire and the interest and task ease questions. Task order was partially randomized; four surveys were created for each condition and the task order in each set manipulated to reduce location bias. Participants were all given the same instructions and tasks. No time constraints for task completion were imposed. Tasks consisted of three parts to ensure the accurate capture of time data relating to time spent on web pages: 1) instructions to find a headline, 2) display of web page screenshot where headline could be found, 3) selection from a list of 4 possible headlines, all drawn from actual news headlines, plus an "I don't know" option. Participants were unable to return to previous pages.

Mechanical Turk: Introduction/Consent, HIT acceptance	
Link to external survey; entry of MTurk worker ID	
Pre-task PANAS	
Pre-task topic interest questions	
8 Non-Salient tasks	8 Salient tasks
Self-report of time spent finding headlines	
Post-task PANAS	
Self-report of affect	
Focused attention scale	
Self-report of focused attention	
Interest and task ease questions	
Monitor size and perception of page loading time questions	
Demographics questionnaire	
Comments	
Return to MTurk to submit HIT	

Table 1 Study protocol

Participants

Participants were drawn from the MTurk population and screened based on location (United States) and HIT approval rate (> 95%), as recommended in the literature [13] and communications with crowd-sourcing experts. The 189 participants (female=65.1%, male=34.4%, prefer not to say=0.5%) were mainly under the age of 41 (79.9%) with the largest group between the ages of 24-29 (32.8%). Participants tended to have a bachelor (39.2%) or some college degree (22.8%). They were primarily employed for a company or organization (42.9%), though there were a number of self-employed (14.8%), students (18.5%), and not employed (20.6%).

Most indicated using a monitor that was 18" or less (73.2%) with the most commonly reported monitor size of 15-16" (30.1%). Regarding page loading time, 23 (12.2%) agreed or strongly agreed that it took a long time for the pages to load. Overall, participants expressed the least amount of interest in Gleeson (N=88, M=1.83, SD=1.06) and the most in Disney (N=189, M=3.33, SD=1.29). Many participants indicated not knowing Worthington (41%) or Gleeson (53%). However, no more than 1% indicated not knowing the other headline topics and 47% of participants agreed/strongly agreed that they follow entertainment news. Unlike a 2010 Pew study that indicates that 47% of Americans get their celebrities or entertainment news online [20], 79% of the participants in our study indicated getting their entertainment news via the internet. No significant differences (p >.05) were found between the two conditions (salient and non-salient) relating to demographics, monitor size, page loading, or entertainment news behaviors.

Analysis

All quantitative analyses were conducted in SPSS. The open-ended responses relating to overall feelings, ease of focus and distraction were analyzed through a process of inductive, thematic coding. The first author developed a coding manual that included definitions and examples of each of the codes. An independent researcher was asked to code 50 responses from each of the three questions using the coding manual. Based on this initial step, code definitions were clarified to improve coding consistency. Both the researcher and the independent researcher then re-coded all three sets of responses independently using the revised coding manual. Kappa was used to determine the reliability of the coding; Kappas > .75 are considered excellent agreement beyond chance [3]. This criteria was met for all three questions (overall feelings = .85); ease of focus = .78; and distraction = .83). The researchers then reached agreement on those responses where their coding conflicted.

RESULTS

The average time for completion of the study was approximately 13 minutes. Of the 200 participants, 96 (48%) were in the salient condition and 104 (52%) were in the non-salient condition. Of these, two were found to have repeated the study after abandoning the study part-way through therefore these were deleted. Those unable to correctly answer at least 6 of the 8 tasks (N=2) were also eliminated. Given the easy nature of the tasks, these participants were not making a sufficient attempt to find the headlines. To the "confident I found the headlines"

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Pre-PAS	28.28	8.94	-								
2. Post-PAS	27.47	9.65	.88**	-							
3. Focused Attention Scale	20.72	6.84	.30**	.44**	-						
4. Time on web pages	3.38	1.92	.13	.16*	.19*	-					
5. Perceived time	5.38	3.36	.14*	.14*	.15*	.53**	-				
6. Easy to locate headlines	4.44	0.81	.22**	.23**	.12*	-.18*	-.11	-			
7. Wanted to click links	3.19	1.40	.24**	.32**	.42**	.22**	.09	.06	-		
8. Interest in webpage content	3.34	1.26	.33**	.42**	.50**	.27**	.10	.05	.79**	-	
9. Wanted to find out more	3.27	1.34	.27**	.35**	.50**	.19**	.05	.12	.78**	.84**	-

Note. $N = 189$. *$p < .05$, **$p < .01$.

Table 2 Summary of Inter-correlations, Means, and Standard Deviations of Study Variables

statement 92.3% responded "agree" or "strongly agree"; no further analysis was performed due to this unanimity.

All missing values were NMAR (not missing at random) and included responses such as "don't know" and "cannot say/not applicable". Six data sets were removed as a result of this NMAR data, leaving *N*=190. The negative affect scale (NAS) of the PANAS was non-normally distributed, showing extreme signs of skew and kurtosis even after log base-10 transformation. As NAS violated the assumption of linearity required for this study's analyses and as our main focus of this study was the PAS rather than the NAS, no further analysis was done using the NAS.

Univariate outliers were examined; a number were identified but their removal did not change the results therefore these data sets were retained. Using Mahalanobis distance five multivariate outliers were identified ($p < .001$); however, Cook's distance values for all but one of the cases fell below the generally accepted cutoff of 1 [23]. This case was deleted, leaving *N*=189 for the analyses reported in this paper. Table 2 contains the descriptive statistics and inter-correlations of the study variables.

There was a relatively even split by condition with 90 (47.6%) in the salient task condition and 99 (52.4%) in the non-salient task condition. Of these, 83.1% ($N = 157$) answered all 8 news headline tasks, 12.7% ($N = 24$) answered 7 of the tasks correctly and 4.2% ($N = 8$) answered 6 of tasks correctly. There were no significant differences by condition relating to task accuracy.

Several themes emerged from the qualitative analysis of the three open-ended questions relating to overall affect, easy to focus, and distractions. Only those themes relating to our hypotheses and exploration of user interests, focused attention, and affect are explored in this paper. Discussion of these themes is interwoven through the discussion of the quantitative results in the following sections. Results are organized according to our hypotheses and exploratory examination of user interests, focused attention and affect. Participants are referred to by condition and number. For example, S130 refers to participant 130 in the *salient* condition; NS21 refers to participant 21 in the *non-salient* condition.

Hypothesis 1: Task performance and saliency

Saliency of task-relevant information leads to greater task performance, both actual and perceived.

Analysis of actual time spent (in minutes) finding the headlines indicates a significant effect for saliency, $t(170.06) = 3.60$, $p < .001$; those in the salient condition ($M = 2.77$, $SD = 1.75$) found the headlines faster than those in the non-salient condition ($M = 3.94$, $SD = 2.68$). As well, there was a significant effect for saliency, $t(182.43) = -2.49$, $p=.01$ with those in the salient task condition finding it easier to locate the headlines ($M = 4.59$, $SD = .69$) than those in the non-salient condition ($M = 4.30$, $SD = .89$). However, does this positive effect on task performance translate to positive feelings and emotions upon task completion?

Hypothesis 2: Affect and saliency

Affect improves when the task-relevant information is salient compared to when it is not salient.

There was no significant effect for saliency between the two groups relating to post positive affect, $t\ (187) = -.35$, $p = .73$. Positive affect, however, significantly decreased from pre- to post-task $t\ (188) = 2.41$, $p = .02$. Pre- and post-PAS scores were examined for each of the two conditions. In the non-salient condition there was a significant difference between positive affect before ($M = 28.3$, $SD = 9.1$) and after ($M = 27.2$, $SD = 10.1$) the tasks were performed, $t\ (98) = 2.25$, $p = .03$, with a decrease in affect post-task representing a small effect size of $r = .23$. This same result was not observed in the salient condition, $t\ (89) = 1.09$, $p = .28$, where there was no change in affect between pre- and post-task affect scores.

Analysis of the "Overall, did you feel positive or negative while completing the news headline tasks?" question revealed no differences by condition, $\chi^2\ (2, N = 189) = 2.25$, $p = .32$. Thematic analysis of responses to the why participants felt positive of negative, or somewhat positive

or negative revealed similar patterns in both the salient task group and non-salient task group responses. Both groups, for example, tended to associate positive feelings with performance, indicating the tasks were easy or relatively easy to perform. One participant wrote: "I felt like I was able to find the headlines quickly and efficiently" (S121) while another noted "I felt positive because I was able to find the headlines quickly" (NS40). Despite the ease of the tasks, another theme to emerge across both conditions was that the web pages were cluttered or there was too much content. Negative feelings emerged because "There was too much at which to look. The page was too busy" (NS29).

Hypothesis 3: Focused attention and saliency

Saliency of task-relevant information leads to higher levels of focused attention.

There was no significant effect for saliency, t (187) = -.09, p = .93 in the focused attention scale scores. Likewise, there were no significant differences in their likelihood to overestimate or underestimate time spent finding the headlines (χ^2 = 3.43, df = 1, p = .06). Condition also had no significant impact on the extent of overestimation (F = .06, df = 1, 149, p = .81), underestimation (F = .002, df = 1, 36, p = .97), or absolute error in estimation (F = .04, df = 1, 187, p = .83).

Responses to the question, "Did you find anything distracting when you were searching for the headlines?" did not differ by saliency condition, χ^2 (2, N = 189) = 5.78, p = .06. However, the open-ended responses relating to distraction suggests that those in the non-salient condition appeared more apt to be distracted by web page format, layout or image saliency. One participant (NS100) wrote: "Some of the pictures and large graphics were distracting when the headline I was looking for was in small type".

This pattern of responses is reflected in the responses to the question "Was it easy to focus on finding the news headlines you were asked to find?" There were significant differences by saliency condition, χ^2 (2, N = 189) = 7.73, p = .02 (Table 3). Due to some cell counts of less than five, a Fisher exact test was performed which confirmed the significance of this result (p < .01). Those in the non-salient condition were more likely to be "on the fence" in their response to this question with more indicating "somewhat" than their salient condition counterparts.

Easy to focus?	Condition		χ^2
	Non-Salient (N=99)	Salient (N=90)	
Yes	81	85	7.73*
Somewhat	17	4	
No	1	1	

Note. * p < .05

Table 3 Cross Tabulation of Easy to Focus and Condition

Responses to the "why?" portion of the easy-to-focus question indicated that saliency was perceived to play a role in whether or not they found it easy to focus on finding the headlines. Those in the salient condition indicated that the location or (more often) the salience of the task headlines made it easy to focus. S14 noted that it was easy to focus "because they were all close to the top and had a picture or were bolded or highlighted" while S74 wrote "Because most of the headlines that I was looking for stood out from the rest".

Exploration: User interests, focused attention and affect

While saliency impacted focused attention and affect, a prominent theme to emerge from the thematic analysis of the open-ended questions was the impact that web page content had on participants' focus. One participant wrote: "My attention was drawn at first to headlines regarding things that I am particularly interested (i.e., the Grey's Anatomy episode blurb). But once I gave myself an attention check, it was easy to find the headlines" (NS190). Others were able to stay focused due to lack of interest in the celebrity news: "I don't really care about celebrity news, so I didn't get distracted very easily" (NS181).

Multiple regression analysis was used to test whether interest in topics, wanting to click on the links, and wanting to find out more were related to focused attention scores (results summarized in Table 4). The results indicated that the three independent variables explained 27% of the variance. However, while interest in content (β = .30, p < 0.01) and wanting to find out more (β = .27, p = 0.03), each contributed significantly to the prediction of focused attention scores, wanting to click on links was not a significant predictor (β = -.03, p = 0.78).

Model	*b*	*SE b*	*β*	*t*	*Sig.*
Constant	11.27	1.23		9.19	.00*
Interest in content	1.64	.67	.30	2.45	.02
Wanted to click on links	-.15	.53	-.03	-.28	.78
Wanted to find out more	1.36	.63	.27	2.17	.03

Note. R^2 = .27, *p < .001.

Table 4 Multiple Regression for Focused Attention Scores

As well as engagement in the web page content, engagement in the task itself appeared to be an integral reason behind open-ended responses. In order to focus, one participant noted that "[…] I made it into a game with myself" (S182). Another who found it easy to focus wrote that "It was kind of like a puzzle trying to find them amid all the other useless information on those pages" (S22). Likewise, the gaming aspect led some to express positive emotions as a result of the experience: "I enjoyed the game of it […]" (NS98); "It's interesting, like a hide and find" (S87). Others expressed neutral feelings relating to their experience noted "It was just a game of finding the headlines, but not that interesting a game" (S169). This

leads to the question: does focused attention, an aspect of user engagement, predict positive affect?

Simple linear regression was used to test whether focused attention scores significantly predicted post-positive affect. Regression results are summarized in Table 5.

Model	*B*	*SE b*	*β*	*T*	*Sig.*
Constant	14.66	2.03		7.24	.00*
Focused Attention	.62	.09	.44	6.66	.00*

Note. $R^2 = .19$, $*p < .001$.

Table 5 Simple Linear Regression for Post Positive Affect

The results of the regression indicated that focused attention scores contributed significantly to the prediction of post-positive affect scores ($\beta = .44$, $p < .001$) and explained 19% of the variance. Pre- positive affect also significantly predicts focused attention ($\beta = .30$, $p < .001$), and explained 10% of the variance.

DISCUSSION

Saliency of task-relevant information leads to better task performance, both actual and perceived (H1)

When the task relevant information is salient, it is found faster, and participants report that it is easier to find the headlines. This hypothesis is thus fully supported.

The connection between task difficulty and time on task has been noted previously [5] and this study confirms that saliency contributes to both subjective measures of task difficulty and actual measures of time. Salient headlines helped participants complete the task more quickly. While the tasks were easy for both groups to complete, those provided with salient headlines perceived the tasks were easier. But does this benefit in task performance translate to a positive affect?

Affect improves when the task-relevant information is salient compared to when it is not salient (H2)

This hypothesis was supported through the results of the pre- and post- positive affect scores, though unsupported through the responses to the overall affect question.

In general, entertainment news web pages are cluttered and have several visual elements; there are a multitude of images and advertisements, headlines and bodies of text on a single page. This is also true of the web pages that our participants were shown; there was a lot of information on the pages. Web pages that are cluttered demand a lot of attention [19] and the stylistic properties of retrieved documents such as presentation of information have been found to arouse both positive and negative emotions [11, 17]. Emotions have also previously been shown to change from positive to negative as task difficulty increases [1]. Saliency of relevant information provided a way to navigate through the page clutter more effectively, making tasks easier for those in the salient condition. This in turn appears to have allowed their emotions to remain positive and stable from pre- to post-task, mitigating a decrease in positive affect. Previous studies [5, 11] examining task type and topic as factors contributing to mood changes before and after online information searching indicated that mood remained stable, however these studies manipulated the tasks rather than the features of the web pages themselves.

Saliency of task-relevant information leads to higher levels of focused attention (H3)

This hypothesis was unsupported through the sum scores of the focused attention scale and estimation of the time spent finding headlines. Focused attention does not increase when the task relevant information is salient nor does a lack of saliency increase distortions of time perception. In contrast, users reported that it was easier to focus when the headlines appeared salient. Why the contradiction?

Recall that responses to the focused attention scale were based on the *experience* of finding news headlines on the web pages. The tasks themselves were easy, short, and involved only a single change in headline saliency between the two groups. The user engagement scale was developed to measure engagement in interactive systems [18]. Perhaps engagement with multiple non-interactive web page screenshots in the entertainment news was not enough to make an impact on perceived focused attention based on saliency alone. Measurement of any potential differences was thwarted possibly in part by our study design. Allowing interaction with web pages, however, would have made it impossible to control saliency to the degree that we did in this study. Due to the lack of interaction with the web page screenshot, the focused attention subscale may have been too abstract. As evidenced in the qualitative analysis, participants reported that it was easier to focus in the salient condition, suggesting that this question helped them think more concretely about what helped them focus.

Like focused attention, distortion of time perception was expected to be greater when the task-based information was salient than when it was not salient. Time distortion or lack of awareness of time occurs when the user is focused on or immersed in an activity [18, 15]. While we expected that saliency would help participants maintain their focus, become engaged with the task and therefore lose track of time, this did not occur. Perhaps it was the case that the task was simply too easy for those looking for salient headlines; a certain amount of challenge is necessary if a user is to become engaged or immersed in their activity and lose themselves, often referred to as the experience of flow [9]. Jennett et al. [9] found that level of immersion in a game was higher than that of immersion in a controlled activity. As both participants in the salient and non-salient conditions were asked to perform tasks that were essentially controlled activities, this level of control may have mitigated against focused attention in both conditions.

Exploration: User interests, focused attention and affect

Our exploratory analysis of the interactions between user interests, focused attention and positive affect suggest a

number of relationships between these both within and outside the remit of saliency.

Relationships between user interests and focused attention
Interest in content predicts focused attention. Similar to other findings [6, 19], our study suggests that when engaged in a goal-based task, users often make a concerted effort to stay focused and block out non-relevant web page content despite efforts to draw their attention. As responses to the follow-up questions suggested, some made a game out of the tasks to maintain their focus. By doing so, they tapped into their attention mechanism, allowing them to control which visual stimuli are processed [19]. By making relevant, goal-oriented information salient, this aids the attention mechanism, allowing users to complete the task-at-hand [19].

Despite the focused task we designed, non-relevant content not only captured attention, but also curiosity, and played a significant role in the perception of distraction during the tasks. This reinforces previous findings suggesting that content and web links of interest to or relevant to the user's goal positively influence focusing of attention on the information-seeking activity; likewise, interesting content and links not relevant to the user's goal have a positive influence on the user's *specific* or *diversive* curiosity in a topic [19]. It seems intuitive, but the degree of interest people have in the content they are encountering has an impact on how engaged they become with this content. A recent quasi-experimental study by O'Brien [17] concluded that the user engagement scale which includes the focused subscale should be expanded to include content as a quality of user engagement due to the emergence of a relationship found between content and the way it is presented on web pages.

Relationship between focused attention and positive affect
Immersion in gaming can sometimes lead to positive affect and other times negative affect [9]. Both negative and positive emotions have also been expressed in relation to reading online news [17]. Our results suggest that focused attention, whether influenced by the non-task-based content or by focused attention to the task itself influences how they feel once they have disengaged. While mood prior to entertainment news tasks influences focused attention, there is a strong relationship between focused attention during the tasks and post-task emotions. The relationship between focused attention and affect therefore merits further investigation.

CONCLUSIONS

This paper presented a study of the impact of saliency of task-relevant information on two engagement metrics, focused attention and affect. Participants were recruited via Amazon Mechanical Turk to respond to an online survey that included information seeking tasks; half were provided with web pages where the task-related headlines were salient while the other half were given web pages where the headlines were not salient. This paper makes three novel contributions:

- This paper detailed the first study showing how saliency impacts metrics associated with user engagement (positive affect and focused attention).
- We showed that when the headlines that users seek appear visually non-salient, users are slow at finding them, report more distraction due to web page features, and show a drop in affect. In comparison, when the headlines appear visually catchy or salient, they find them faster, report that it is easy to focus, and maintain positive affect. This shows that saliency is helpful in task performance, focusing/avoiding distraction and in maintaining positive affect.
- By adapting the focused attention subscale from the online shopping domain [16] to the entertainment news domain, we studied how saliency impacts focused attention. Although users reported that it was easier to focus in the salient condition, we found no significant improvement in the focused attention subscale or differences in perceived time spent on tasks. Instead, we find that user interest in web page content is a good predictor of focused attention, which in turn is a good predictor of positive affect.

Our results ultimately suggest that the interaction of saliency, focused attention, and affect, together with user interest, is complex.

Examining saliency and engagement simultaneously in this study was a challenge. Salient features of a web page draw user attention within the first few seconds of visiting a page, but engagement happens after, sometimes long after, that initial exposure. Moreover, the impact that the manipulation of our perceptual environment has on our feelings and emotions is difficult to measure. Nonetheless, the use of multiple measures helped us gain a richer understanding of the complex relationships between saliency, focused attention, affect and user interest.

Performing this study online with no control over screen size, internet connection speed, or other variables we would normally control for in a laboratory setting presented further challenges. However, while the use of Amazon Mechanical Turk as a platform for online recruitment took much of the environmental control away that we could have maintained in a lab setting, it also allowed for more rapid testing and a larger and more diverse participant base than the average user study. In addition, the Mechanical Turk Workers made excellent participants, contributing through their post-study comments suggestions for study improvements. Responses to open-ended questions were equally thoughtful and contributed significantly to the interpretation of our results.

While we were limited in the present study to working with static web pages, future studies should allow participants to

engage more naturally with the website content. As well, we were limited in this study to static forms of salience, the manipulation of color, images, and font size. Future studies should explore the impact that common dynamic web page features such as animation or simple flickering may have on affect and focused attention. Findings would be of interest to any website owner, entertainment-oriented or other, interested in not only in catching users' attention, but also maintaining it.

We did not control for interest in this study, but future studies could incorporate this into the design by gathering users' interests on different topics in advance (e.g., sports, finance, entertainment, travel, fashion). A more naturalistic study could subsequently be conducted. Salience of information that the participant is known to enjoy or have an interest in could be manipulated to test how saliency, user interests, affect and focused attention, and perhaps other aspects of user engagement, interact. We speculate that increasing the salience of information of personal interest to users might improve focused attention and create positive affect, though this needs to be tested rigorously in future work. Rather than imposing a goal-oriented, externally-imposed task, facilitating more natural interactions will broaden our understanding of the impact of saliency on users.

ACKNOWLEDGMENTS

We would like to thank Diego Sáez for his support in implementing the online study and Ioannis Arapakis for assistance with the qualitative data analysis. This study was completed while the first author was on an internship at Yahoo! Research in Barcelona, Spain.

REFERENCES

1. I. Arapakis, J.M. Jose & P.D. Gray. Affective feedback: An investigation into the role of emotions in the information seeking process. In *Proc. SIGIR2008,* ACM press (2008), 395-402.
2. S. Attfield, G. Kazai, M. Lalmas & B. Piwowarski. Towards a science of user engagement (Position Paper), WSDM Workshop on User Modelling for Web Applications (2011).
3. M. Banerjee, M. Capozzoli, L. McSweeney, & D. Sinha. Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics, 27,* 1 (1999).
4. R. Carmi & L. Itti. Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research, 46,* 26 (2006).
5. J. Gwizdka & I. Lopatovska. The role of subjective factors in the information search process. *JASIST, 60,* 12 (2009).
6. M.H. Huang. Designing website attributes to induce experiential encounters. *Computers in Human Behavior, 19*, 4 (2003).
7. L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 10-12 (2000), 1489-1506.
8. M. Jennings. Theory and models for creating engaging and immersive e-commerce websites. In *Proc. SIGCPR 2000*, ACM Press (2000).
9. C. Jennett, A.L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs & A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 9 (2008), 641-661.
10. A. Kittur, E.H. Chi & B. Suh. *Crowdsourcing user studies with Mechanical Turk*. In *Proc. SIGCHI* 2008, ACM Press (2008).
11. I. Lopatovska. Searching for good mood: Examining relationships between search task and mood. *American Society for Information Science and Technology, 46,* 1 (2009), 1-13.
12. C.C. Marshall & F.M. Shipman. The ownership and reuse of visual media. In *Proc. JCDL 2011*, ACM Press (2011).
13. W. Mason & S. Suri. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods* (2010), 1-23.
14. V. Navalpakkam & L. Itti. Search goal tunes visual features optimally. *Neuron, 53,* 4 (2007), 605-617.
15. T.P. Novak, D.L. Hoffman, & Y. Yiu-Fai. Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Marketing Science, 19,* 1 (2000), 22-42.
16. H.L. O'Brien. *Defining and Measuring Engagement in User Experiences with Technology.* PhD Thesis, 2008.
17. H.L. O'Brien. Exploring user engagement in online news interactions. In *Proc. ASIST 2011* (forthcoming).
18. H. O'Brien & E. Toms. The development and evaluation of a survey to measure user engagement. *JASIST 61*, 1 (2010).
19. S. Pace. A grounded theory of the flow experiences of Web users. *International Journal of Human-Computer Studies, 60,* 3 (2004), 327-363.
20. K. Purcell, L. Rainie, A. Mitchell, T. Rosenstiel, & K. Olmstead. *Understanding the Participatory News Consumer*. Pew Internet & American Life Project (2010).
21. K. Overbeeke, T. Djajadiningrat, C. Hummels, S. Wensveen & J. Frens. Let's make things engaging. *Funology*. Kluwer, 2003.
22. R. Peters & L. Itti. Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception (TAP), 5*, 2 (2008).
23. B.G. Tabachnick & L.S. Fidell. *Using Multivariate Statistics* (4th ed.). Boston, MA: Allyn and Bacon, 2001.
24. A. Treisman & G. Gelade. A feature-integration theory of attention. *Cognitive Psychology, 12,* 1 (1980).
25. D. Watson, L.A. Clark & A. Tellegen. Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 47* (1988).
26. J. Wolfe & T. Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 6 (2004), 495-501.