Exploring Website Gist Through Rapid Serial Visual Presentation

Justin W. Owens
Google, Inc.

Barbara S. Chaparro
Embry Riddle Aeronautical University

Evan M. Palmer
San Jose State University, evan.palmer@sjsu.edu

Exploring Website Gist Through Rapid Serial Visual Presentation

Justin W. Owens[1,2], Barbara S. Chaparro[1,3], Evan M. Palmer[1,4]

[1] Department of Psychology, Wichita State University

[2] Google, Inc.

[3] Department of Human Factors and Behavioral Neurobiology, Embry Riddle Aeronautical University

[4] Department of Psychology, San José State University

SUPPLEMENTAL MATERIALS

Color Energy and Picture Frequency Analyses of Website Stimulus Set

The following analyses represent an initial attempt at understanding the low-level image differences between the website categories used in the study. Informal inspection of the stimulus set suggests some systematic differences between website categories, such as search stimuli having a lot of white space and blue text and shopping stimuli having many pictures of people and other objects. To more rigorously formalize these differences, we performed some simple analyses of the stimulus set. We are also posting our stimulus set for the research community so that others may explore these differences with more sophisticated analyses or use the stimuli in further research.

Stimulus Set Analyses. To delve deeper into properties of our stimulus set that may have aided detection, we quantified several low-level characteristics of our stimuli. Our first analysis examines the relative number of white, black, red, green, and blue pixels in the different website categories. Our second analysis focuses on the number of pictures in our stimuli, and whether they depict people or other objects.

Color Analysis. For each image in the stimulus set, every pixel has a value between 0 and 255 for the red, green, and blue channels. Using these values, we calculated the number of predominately white pixels (red+green+blue values >= 245*3), black pixels (red+green+blue values <= 10*3), red (red values > green+blue and > 128), green (green values > red+blue and > 128), and blue (blue values > red+green and > 128) pixels for all 1104 images in our stimulus set. We then analyzed these data using a 4 (website category) x 5 (color) mixed ANOVA, with image category as a between-subjects variable and image values as a repeated measures variable, and employing Greenhouse-Geisser corrections for sphericity violations where appropriate. Planned comparisons were conducted using two-tailed t-tests with Bonferroni corrections to guard against type I error.

Figure 1 depicts the proportion of color energy in our stimulus set as a function of website category. Analyses of the color properties of the stimuli revealed a main effect of website stimulus category, $F(3,1100) = 103.00$, $p < .001$, $\eta^2 = 0.22$, a main effect of color, $F(2.30,2530.96) = 3324.09$, $p < .001$, $\eta^2 = .71$, and a website category x color interaction, $F(6.93,2530.96) = 79.81$, $p < .001$, $\eta^2 = .051$. Focusing on the interaction, we analyzed the relative amounts of color energy as a function of website category. Search websites had significantly more white energy than all other website categories, all $t(550) \geq 9.67$, all $p \leq .001$, all $d \geq .86$, followed by shopping, all $t(550) \geq 3.06$, all $p \leq .001$, all $d \geq .26$, and news, $t(550) =$

7.14, all $p \leq .001$, all $d \geq .61$. Social websites had more black pixels than all other categories, all $t(550) \geq 4.18$, all $p < .001$, all $d \geq .36$, followed by news, all $t(550) \geq 5.61$, all $p < .001$, all $d \geq .48$, then shopping and finally search, $t(550) = 5.75$, $p < .001$, $d = .49$. Shopping websites had more red pixels than all other website categories, all $t(550) \geq 3.26$, all $p \leq .001$, all $d \geq .28$, followed by news, all $t(550) \geq 4.1$, all $p \leq .001$, all $d \geq .35$, then social, which had more red pixels than search, $t(550) = 6.28$, $p < .001$, $d = .54$. In terms of green pixels, social and search websites had higher values than all other categories, all $t(550) \geq 3.92$, all $p < .001$, all $d \geq .33$, but were not significantly different from each other, $t(550) = .58$, $p = .56$, $d = .049$, n.s. The number of green pixels in news and shopping website stimuli also did not differ, $t(550) = .38$, $p = .70$, $d = .033$, n.s. The pattern of results for blue pixels was more complicated, with search websites having more blue pixels than news, $t(550) = 2.65$, $p = .008$, $d = .23$, or shopping, $t(550) = 3.93$, $p < .001$, $d = .34$, but not social, $p = 39$, n.s. Social websites had more blue pixels than shopping websites, $t(550) = 2.76$, $p = .006$, $d = .24$, but not news, $p = .14$, n.s. Finally, the news and shopping sites did not have a statistically different number of blue pixels, $p = .054$, n.s.

To summarize, some color energy patterns varied in reliable ways for the website categories, but not necessarily in ways that would predict the behavioral data. Shopping and search websites were the most easily detected in Experiment 1. Shopping websites had the most red pixels, while search websites had the most white pixels and were tied for the most blue pixels with social websites. Social websites had the most black pixels and were tied with search for the most green pixels. Finally, news websites tended to have a moderate number of white, red, and blue pixels compared to the other categories. It is possible that participants could have performed best in the shopping and search categories by looking for red and white/blue colors in our stimulus set, respectively, but whether participants rely on such color information to perform the task is an open question worthy of future research.

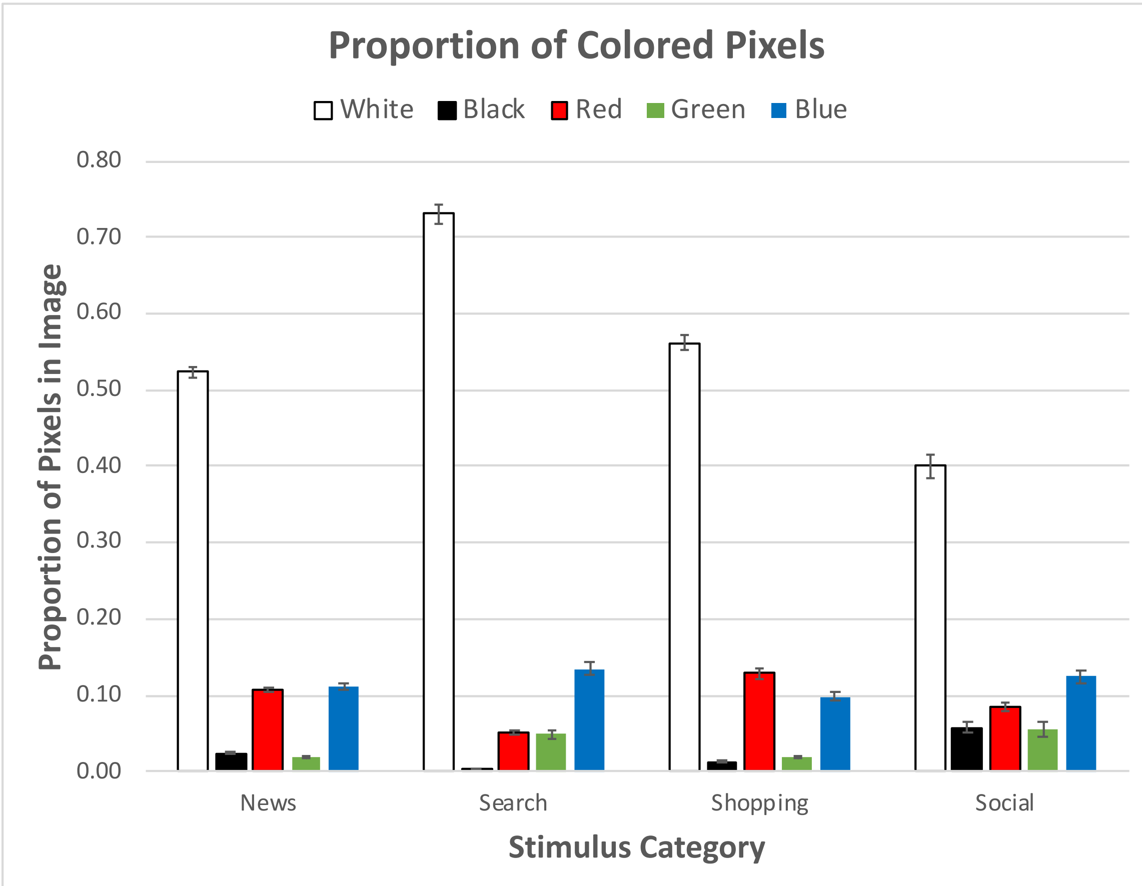

Figure 1. Percentage of colored pixels by website stimulus category. Error bars depict standard error of the mean.

Picture Analysis. We counted the number of pictures in each webpage stimulus, classifying them as either pictures featuring people (images with any human body part visible) or other pictures (images with no human body parts visible, excluding text-only images; Figure 2). We then submitted these data to a 4 (website category) x 2 (image type) mixed ANOVA with website category as a between-subjects variable and image type as a within-subjects variable.

Planned comparisons were conducted using two-tailed t-tests with Bonferroni corrections to guard against type I error.

The analyses detected a main effect of image category, $F(3,1100) = 176.80$, $p < .001$, $\eta^2 = .33$, and planned comparisons revealed that social media websites had the most pictures overall, all $t(550) \geq 6.64$, all $p < .001$, all $d \geq .74$, followed by news and shopping, which did not significantly differ, $p = .011$, n.s. Search webpage stimuli had the fewest number of images per page, compared to the other three stimulus categories, all $t(550) \geq 17.50$, all $p < .001$, all $d \geq 1.49$.

There was no main effect of picture type, meaning that overall the number of pictures with and without people was roughly proportionate across website categories. However, there was a significant interaction of picture type by website category, $F(3,1100) = 138.94$, $p < .001$, $\eta^2 = .27$, indicating that the number of pictures of people versus other objects differed by website category. Planned comparisons revealed that there were significantly more pictures of people than other objects in the news, $t(275) = 4.49$, $p < .001$, $d = .27$, and social website stimuli, $t(275) = 18.87$, $p < .001$, $d = 1.14$, but more pictures of objects than people in the search, $t(275) = 5.26$, $p < .001$, $d = .32$, and shopping stimuli, $t(275) = 15.89$, $p < .001$, $d = .96$.

Based on this analysis, participants may have performed best in the shopping and search categories if they detected stimuli with a higher proportion of non-human than human pictures. On the other hand, they should have been able to recognize news and social media/blog stimuli by the relatively high number of pictures featuring humans. Both news and search have a more equal (though still significantly different) ratio of human vs. non-human pictures, with news websites having far more images, on average, than search websites. Again, these analyses are suggestive but more focused work exploring these variables needs to be conducted before we can come to any strong conclusions about the sorts of low-level stimulus information participants may be using to perform this task.

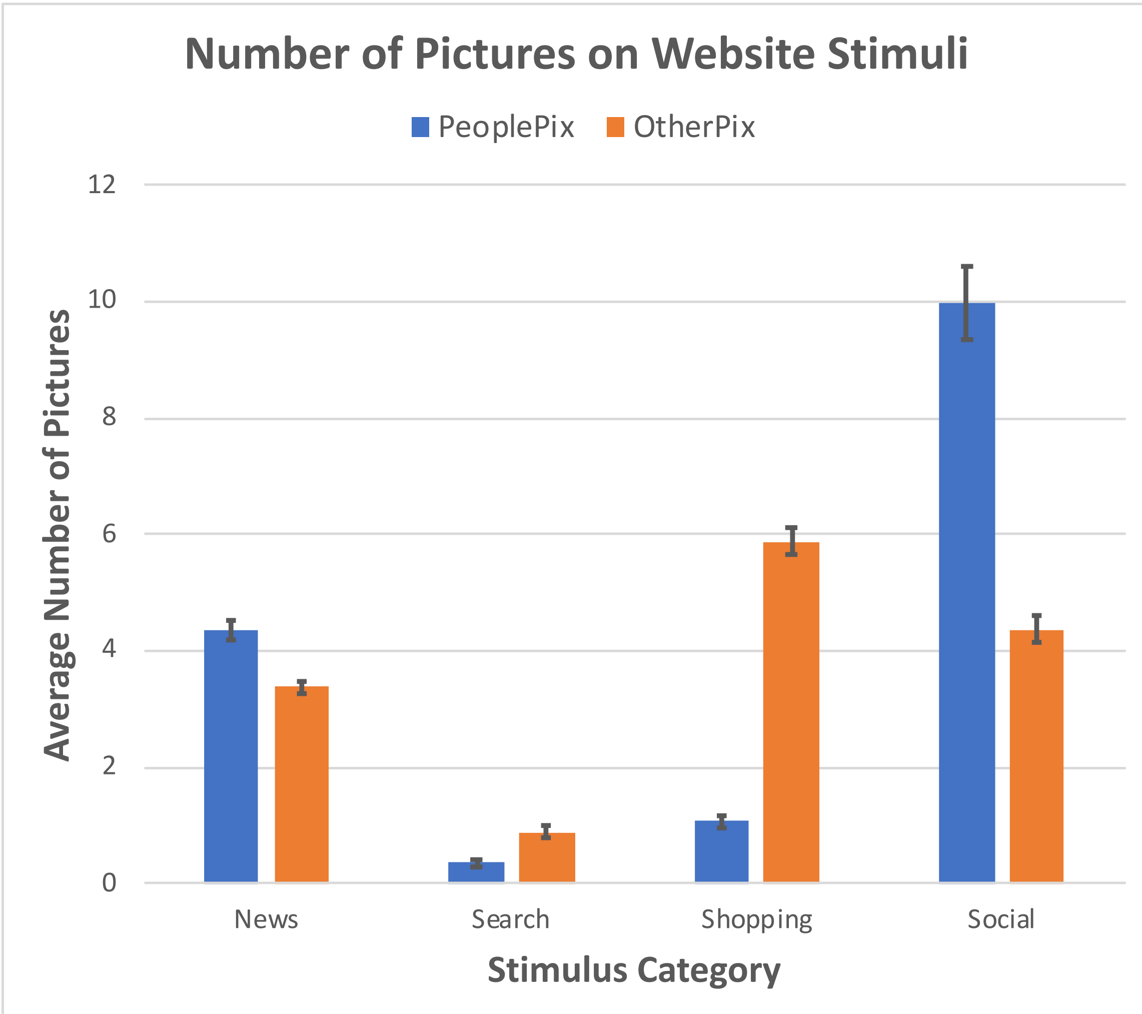

Figure 2. Average number and type of pictures as a function of stimulus category. Error bars reflect standard error of the mean.