The Use of Change Detection as a Method of Objectively Evaluating Labels

By Laura Bix,1* Wataru Kosugi,2 Nora Bello,3 Raghav Sundar1 and Mark Becker4

1School of Packaging, Michigan State University, East Lansing, MI 48824, USA
2Research & Development for Beverages, Asahi Breweries, Ibaraki 302-0106, Japan
3Department of Animal Science, Michigan State University, East Lansing, MI 48824, USA
4Department of Psychology, Michigan State University, East Lansing, MI 48824, USA

SUMMARY

This study introduces the use of change detection, a technique used in cognitive psychology to measure attentional scan paths, as a way to objectively evaluate the prominence of varied label elements. There are two major objectives related to this work: (a) to develop change detection software and methodology for label use; and (b) to compare the relative prominence of different label elements on a beverage container. Six label elements (i.e. treatments) were analysed, namely: the manufacturer name, the product name and a warning dot with text in three colours. Study results suggest that experimental set-up can significantly impact results, specifically the position of the change ($p = 0.0078$) and the order of appearance ($p = 0.069$). This is not only important from an experimental design perspective, but also could lead to insights regarding the attentive behaviours of people as they purchase, select and use products. With regard to the elements of the labels tested, we identified a significant difference on time to detect a change ($p < 0.0001$). Time required for the manufacturer’s name, Asahi Breweries, was significantly longer than for any of the other label elements ($p < 0.0001$). Pairwise comparisons indicated that for the warning dot, red text was located marginally faster than the warning printed in black ($p = 0.0566$). Change detection offers the promise to objectively evaluate the relative prominence of a label (or a scene) but is quicker and cheaper than other methods, such as eye tracking, that are currently utilized for this purpose. Copyright © 2010 John Wiley & Sons, Ltd.

Received 18 November 2009; Revised 5 April 2010; Accepted 12 April 2010

KEY WORDS: prominent; conspicuous; labels; labelling; beverage labels

INTRODUCTION

Most people believe that our perceptual systems are passive agents that veridically transfer the external world into an internal representation. Under this view, all that is required for visual perception to occur is for a visually resolvable stimulus to impinge upon the eye. In reality, perception is an active process that depends critically on limited attentional resources.

The critical role of that attention plays in awareness is evidenced by numerous failures of visual awareness that occur when attentional capacity is scarce.1–4 For instance, participants who were given an attentionally demanding task failed to notice when an unanticipated person wearing a gorilla suit walked through the display. This failure occurred even though the gorilla was highly visible and easily...
detected when the attentionally demanding task was removed.\textsuperscript{1} This example and many others\textsuperscript{2–4} suggest that conscious visual awareness depends critically on attention.

This critical linkage between attention and awareness has implication for the design and evaluation of effective labels. It suggests that effective labels must not only be visually resolvable, but also must attract attention. This conclusion suggests that understanding how label design influences the attentional prioritization of different aspects of a label may be important to the development of optimal labels. Moreover, it suggests that methods used in visual cognition to investigate the linkage between attention and awareness may be applicable to the study of effective labelling.

One such method is a change detection method. Using a variety of change detection methods, researchers have demonstrated that the successful detection of a visual change depends on attention: without attention, participants experience ‘change blindness’, the inability to detect large changes in a scene. Change blindness occurs when two disparate scenes are separated by a blank screen (the flicker change detection technique),\textsuperscript{1–5} eye blinks,\textsuperscript{6} saccades,\textsuperscript{7–9} camera cuts\textsuperscript{10} or occluding events.\textsuperscript{11} It can also be produced when multiple ‘mud-splashes’ hit the screen simultaneous with the change,\textsuperscript{12} or when changes occur extremely slowly.\textsuperscript{13} All of these techniques have one thing in common: the change fails to produce a sudden and unique visual transient that would otherwise capture attention to the location of the change. Thus, these findings highlight that attention to a specific aspect of a scene is required to detect changes to that aspect of the scene.

Here we apply the flicker change detection test to the evaluation of product labels in an attempt to objectively evaluate the attentional prioritization of the different aspects of a label. During the flicker task, participants view an original scene and a modified scene that alternate repeatedly, separated by a brief, blank display. The participant’s task is to detect the change between the original and the modified scene as quickly as possible.\textsuperscript{3,14}

There is substantial evidence from the change detection literature that this type of change detection requires one to attend to and encode the change object into a more durable form of working memory.\textsuperscript{3,15,16} Given this reliance on attention, the time required to detect a change is a good indication of the time when attention first selected that object. Faster detection of the change implies early attention to the changed property, which can also provide insights about the viewer’s ‘attention scan paths’.\textsuperscript{3,14} As such, change detection has been used to measure the locus of attention (a combination of perception and cognition) much as eye tracking measures the fixation of the eye (i.e. perception).\textsuperscript{3,14}

For this study, we introduce the use of the flicker method of change detection as a technique for the objective evaluation of the prominence of varied label elements that appear on a Japanese product. The use of an unfamiliar product with textual information in a foreign language (test subjects were American) meant that previous experience or product familiarity was unlikely to interfere with results. Additionally, the relative conspicuousness of the text that appeared on a warning dot on the label was of interest to the manufacturer, Asahi Breweries (Tokyo, Japan). The purpose of the tested warning dot is to clearly indicate the presence of alcohol in the beverage; this is an important factor, as the vibrant packaging of the product and the appearance of fruits and vegetables on the principal display panel may lead to confusion with non-alcoholic products.

As such, the study evaluated the relative conspicuousness of the following:

- Presence/absence of the warning dot.
- Three different colours of text within the warning dot.
- The product name, ‘Vegesh’.
- The manufacturer name, ‘Asahi’ (see Figure 2).

**OBJECTIVES**

The objectives of this study were to develop software that effectively implements the ‘flicker task’ so that it could be used to objectively evaluate the prominence of varied label elements, and to compare the relative prominence of six different label elements: manufacturer name, product name and a warning dot with text in three colours using the newly developed software, with particular focus on the effect of the colour of the warning text on noticeability.
MATERIALS AND METHODS

Test participants
A convenience sample consisting of 46 student participants (30 male and 16 female) was tested using procedures approved under IRB 09-099. Participants were at least 18 years or older and were screened for a history of seizures. Additionally, they could not be legally blind. Upon arrival, participants were assigned a subject number and were characterized by demographic information that included gender, age, self-reported ethnicity, educational background and language proficiency. Additionally, their visual acuity was tested using a Dow Corning Ophthalmics (Norfolk, VA) near-point visual acuity card, and their ability to see colour was characterized using pseudo Iso-chromatic plates (Richmond Products, Albuquerque, NM, USA). Participants were then subjected to the ‘flicker test’, which consisted of three groups of change detection tasks; each group was separated by a brief period of rest (see Figure 1).

Treatments
Six label elements on a Vegesh can (Asahi Breweries, Tokyo, Japan) were of interest and were defined as ‘treatments’ for flicker testing (see Figure 2).

Flicker testing
For the flicker testing, subjects were comfortably seated in front of a Sony Vaio (Tokyo, Japan) laptop with an Intel Centrino 1.86 GHz processor. The laptop has 1 Gb of RAM and a 15" screen with X-Brite technology. It ran a flicker programme created with Adobe Flash that was developed by the research team. During a trial, a standard image (240 ms) continuously alternated with a test image (240 ms), with a brief, gray screen (80 ms) separating the presentation of the images. This method and timings were based on the methodology used by Rensink et al. The sequence of images, standard–blank–test–blank, looped continuously until the participant responded or the trial timed out at 1 min (see Figure 3).

Each standard image was comprised of four cans, and its corresponding test image was identical to the standard image except that a single aspect of one of the cans was altered (see Figures 2 and 3). There were six changes of interest to the research team (see Figure 2).

During change 1, an image of the standard can alternated with a test can image that did not have the manufacturer’s name present. The second change was similar, but the entire name of the product...
disappeared as the images alternated, while the manufacturer’s name was present in both images. The last four changes represented changes to the alcohol warning. During the first of these (labelled as change 3 in Figure 2), the warning text alternated between black and red; change 6 was similar, except yellow text alternated with black. In change 5, which tested the black warning, black text disappeared as the test can and standard can images alternated. Change 4 marked the disappearance of the warning dot.

Each participant was subjected to three ‘groups’ of change detection tasks. Each group consisted of five pairs of images, two of which were test image pairs, and the remaining three were dummy image pairs. A total of 15 responses were recorded on each subject (i.e. three groups × five image pairs per set). The order of appearance of the two test image pairs and the three dummy image pairs within the group was randomized, as was the order of appearance of the six treatments of interest.

Figure 2. The six test treatments (changes) to the Vegesh can. Each test pair and dummy pair consisted of a standard and an altered picture, with the alteration appearing on only one of four cans.

Figure 3. The flicker task.
Test image pairs were analysed for time to detect the change. Dummy image pairs were added so that participants did not become habituated to the test pairs. However, dummy image pairs were not of direct interest to the research team and, therefore, were not analysed. (see Figure 4 for dummy pairs).

Subjects were instructed that two images would alternate on the computer screen, that one of the images was altered in a single location and that we were trying to see how quickly they could find this change. They were told to stop the test as soon as they detected the change by pressing the space bar on the computer, and then to identify the change by clicking on the area of the label where the change was displayed. In the event that the change was incorrectly identified, the pair was retested, and the times of all trials for the treatment were summed. Time to detect the change was recorded as the dependent variable with the intention of evaluating the noticeability of the six test treatments of interest (see Figure 2). Researchers timed each trial, and the trial was ended if the subject took longer
than 1 min to find the change with 60 s recorded as the time. The binary result (detect yes/no) was also coded as ‘failed to detect’. This was the case for only three data points within the six changes of interest.

Statistical analysis

The response variable, ‘time to detect change’, recorded in seconds, was modelled using a general linear mixed model fitted with the MIXED procedure of SAS software for statistical analysis (Version 9.2, SAS Institute Inc., Cary, NC, USA). The statistical model included the fixed effects of treatment, gender, their two-way interaction, position, group, order of presentation in the group and their two-way interaction. A random effect of subject was also fitted in order to accommodate the split-plot experimental design by which the subject acted as the experimental unit for gender and as the blocking factor for the image pairs. In order to appropriately meet model assumptions, the response variable was log transformed. Residual plots were evaluated, and assumptions were considered to be appropriately met. Least square means estimates and 95% confidence intervals are presented in the original scale. Pairwise comparisons were conducted using Tukey-Kramer’s adjustment to prevent inflation of Type I error rates.

RESULTS AND DISCUSSION

Usable data were obtained from a total of 46 participants, 30 men and 16 women who were between the ages of 18 and 32 years. These 46 subjects each participated in 15 trials for a total of 690 observations. Of the 690 observations, 276 (46 subjects × 6 changes) were of interest to the research team. From the 276 observations of interest, three changes were not detected within the 1 min time frame; two men and one woman failed to see the change in manufacturer name (see Figure 2). This occurred twice when the change was located at position A and once when it was in position B (see Figure 3). As such, 3 out of the 46 observations were recorded and analysed as 60 s.

The statistical analysis identified a significant effect of treatment (see Figure 2) on the time required to detect a change ($p < 0.0001$). Pairwise comparisons that utilized a Tukey-Kramer test indicated that the time required to detect the change in the manufacturer’s name, Asahi Breweries, was significantly longer than for any of the other label elements ($p < 0.0001$), and that changes to the red warning text were located marginally faster than those printed in black ($p = 0.0566$).

The estimated least square mean time to detect a change (expressed in seconds) for each treatment is presented in Figure 5. Letters a through c are used to represent statistical comparisons at $\alpha = 0.05$. For instance, the manufacturer’s name, Asahi, took significantly longer to detect than any other

Figure 5. Estimated least square means of time to detect a change (in seconds, by treatment) with standard error bars. Letters indicate statistical significance at $\alpha = 0.05$. 

element of the label that was evaluated, while no evidence for difference was indicated when the black warning was compared to any element other than the red warning.

Position of the label element change was also indicated to affect the time to detection \((p = 0.0078)\). As with the treatment effect discussed previously, pairwise comparisons were conducted on the position of change (see Figure 6). Changes that occurred in position B took significantly less time to detect than changes that occurred in position A (as indicated by the lowercase letters a and b). However, no difference was indicated between the time required to detect changes at position B when it was compared with C or D, or in A when compared with C or D. This may suggest an order of scanning the image, where subjects may have started from position B, moved right and then back to the far left as they searched the image.

There was also marginal evidence for an effect of group (see Figure 1) on time to change detection \((p = 0.0538)\). When pairwise comparisons were made, subjects detected set 2 marginally faster than set 1 \((p = 0.069)\). This likely occurred as subjects became more familiar with the process as testing progressed. By including the fixed effect of group in our statistical modelling, comparisons between treatments of interest were naturally adjusted for differences between groups. Time to change detection in group 3 was intermediate to groups 1 and 2, and not significantly different from either \((p > 0.11\) in both cases) (see Figure 7). Within groups, we found no evidence for an effect of order of presentation of the stimulus pair on time to change detection \((p = 0.37)\). These results validate our careful randomization of order of presentation of the stimulus pairs.
CONCLUSIONS

We have applied change detection, a tool that is widely used to quantify fundamental behaviours related to perception and cognition, in an attempt to quantify the prominence and conspicuousness of various label elements for Vegesh, an alcoholic beverage from Asahi Breweries. Results suggest that there is a difference in the noticeability of six label elements of interest, as measured by the time to detect a change. Evidence exists for a significant effect of label element (manufacturer name, warning dot and warning in three colours of text and product name) on time to detect ($p < 0.0001$). Pairwise comparisons indicate that the manufacturer’s name took longer to detect than all other elements considered, and that the warning dot with red text was located faster than the warning dot that utilized black text.

Study results also suggest that the experimental set-up of this type of test can have a significant impact on the results, specifically the position ($p = 0.0078$) of the change and group ($p = 0.069$). This is not only important from a design of experiments perspective, but also could lead to insights regarding the attentive behaviours of people as they purchase, select and use products. Further research in this area is recommended.

Change detection is a quantitative tool that is less technical and cheaper to set up, conduct and analyse than eye tracking. As such, it has the potential to be tremendously impactful for both fundamental and applied research in packaging and labelling.

From a fundamental perspective, this tool can be applied in experiments that investigate the search behaviours of people in varied contexts. This could illuminate how consumers search the shelves, potentially impacting sales, or how people search for specific information within a given context (e.g. a nurse searching for the correct drug, with the appropriate route of administration, in an appropriate dosage). Applied research techniques include the ability to quickly and easily evaluate the relative prominence and conspicuousness of varied design elements.

ACKNOWLEDGEMENTS

The authors wish to thank Asahi Breweries for supporting the work of Wataru Kosugi for 1 year at the HUB laboratory at Michigan State University. We would also like to acknowledge Manuel Colunga, who reviewed numerous drafts of this document, significantly improving the figures and the methodology portion, in particular. We would also like to thank several HUB Team researchers: Carly DeHenau and Chris Steckler for their work on the project and the many participants that made this research possible.

REFERENCES

3. Rensink RA, O’Regan JK, Clark JJ. To see or not to see: the need for attention to perceive changes in scenes. Psychological Science 1997; 8(5): 368–373.