Consumer Attention to an Over-the-counter Warning in Four Different Styles of Design

By Abhishek Gawasane, Laura Bix,* Javier de la Fuente, Raghav Prashant Sundar and Timothy J. Smith

School of Packaging, Michigan State University, East Lansing, MI 48824, USA

The study consisted of three objectives: (a) to test the relative prominence and conspicuousness of a warning required by US law to be conspicuous; (b) to explore whether or not the conspicuousness of the said warning can be enhanced graphically; and (c) to develop preliminary data for power analysis that would guide decisions related to sample size in future studies.

Seventeen subjects viewed four over-the-counter drug packages (each with a different style of warning) along with five other products while wearing an eye tracking device. Four styles of warning were used on the over-the-counter drug packages: no outline and no fill, outline and no fill, outline and fill, and outline and fill. The surface area and the placement of the warnings were held constant across all four designs and were consistent with those on commercially available products. Collected data were broken into five zones: warning, brand name, strength, product benefit and net weight.

Despite the fact that US law requires it to be conspicuous, the tested warning was significantly less noticeable than the brand name (objective one) for all dependent variables analyzed ($\alpha = 0.05$). No significant difference was indicated for the varied warning designs (objective two). This could be because not much can be done to enhance prominence when constrained to the limited space that is typically used for such warnings or because of the limited sample size. Power calculations suggest that a sample size of nearly 200 subjects would be required to detect a 2.5 s mean difference at 80% confidence (objective three).

KEY WORDS: conspicuous; over-the-counter; labels; warning; consumer

INTRODUCTION

Risk is pervasive. Whether we are at home, at work or travelling from one place to another, risks are involved. Warnings are intended to help reduce or eliminate risks. However, for a warning to be effective, there must be exposure to the warning, attention to the warning, active processing of the warning, comprehension and agreement with the warning, storage in memory (as well as search and retrieval), response selection and response performance. Failures at any of these stages have the potential to reduce a warning’s effectiveness.1–4

Recent years have seen an increase in warning-related research. Researchers attribute this to varied reasons; the three main reasons are as follows:5

• Increasing healthcare costs (in the USA and around the world).
• Increasing importance of warnings in litigation and court cases.
• Increasing emphasis on warnings by organizations such as the US Environmental Protection Agency and the US Occupational Safety and Health Administration.

---

*Correspondence to: L. Bix, School of Packaging, Michigan State University, East Lansing, MI 48824, USA. E-mail: bixlaura@msu.edu
Warnings are frequently classified into three large spheres. Those that

- Prevent dangerous practices, such as smoking in gas stations or being exposed to sunlight while taking certain drugs.
- Catalyse behavioural change; a warning suggesting protective gear is one such example.
- Enable informed decisions. For example, a warning that indicates the risks associated with alcohol consumption during pregnancy or statutory cigarette warnings.6

The model. We frame our study within the context of an information processing model. Information processing models focus on the stages that warnings pass through while being transmitted from the environment to the user. Such models take into account multiple facets that have the potential to influence the success or failure of the message.6 Information processing theories often mention that the information passes through stages in serial fashion and that failure at any stage will lead to failure of subsequent steps.

A commonly used theory of information processing indicates that the following four steps must be completed for the warning to be effective.2–4,7,6

1. The information must be noticed.
2. It must be encoded into memory. For this to happen, the message must be gathered through perceptual systems and converted from an external representation into an internal one.
3. The encoded message must then be comprehended.
4. It must, finally, move the reader to the appropriate action.

Because of the sequential and serialized nature implicit in information processing, many pieces of legislation and regulation mandate that certain informational items appear ‘conspicuously’ or ‘prominently’ on package labels (see Table 1). Without sufficient prominence, an item is not noticeable, precluding the ability of the consumer to proceed with further processing. Although a limited number of official documents do prescriptively address things like minimum type size and make vague statements like ‘sufficient contrast’, the definitions of ‘prominence’ and ‘conspicuousness’ are vague throughout the mandates.

For this study, we investigated the relative prominence of the warning, ‘This package for households without young children’ (stage one of the information processing model). The Poison Prevention Packaging Act of 1970, 15 U.S.C. pages 1471–1476, requires child resistant closures on most hazardous household substances; this includes drugs as defined in Section 201 of the Federal Food, Drug and Cosmetic Act, 21 U.S.C. page 321.8 One size of over-the-counter (OTC) product may be exempted from the requirement for a child resistant feature, provided it is ‘conspicuously labelled’ with the aforementioned warning. Yet previous research suggests that across several OTC products containing acetaminophen, this warning was consistent in both

- Its design (most were present in the lower left hand corner of the package and of consistent size).
- Its inability to garner attention as measured by eye tracking.9

This led us to wonder ‘if constrained by available space on the principal display panel (PDP), is there a way that warnings could be graphically enhanced to garner attention?’ In other words, if manufacturers are limited to a small space in the lower left hand corner of these packages, is there anything that he carried out to make the warning more noticeable? Holding the warning placement to the lower left hand corner of the PDP and limiting the allowable surface area to a 43 mm × by 4 mm rectangular space (172 mm²), an area and placement based on the commercial products we used for previous study,9 we investigated the effects of two design elements, outline and fill, on the noticeability of the warning as measured by eye tracking.

This is in accordance with the ‘theory of pre-attentive processing’, described by Wogalter et al. as follows:

‘Pre-attentive processing’ suggests that information (e.g. a warning), if designed properly with respect to the background, can ‘pop out’ and attract attention without conscious control on the part of the individual. For example, if an individual is shown a random configuration of 100 equal size dots, a single red dot embedded in 99 blue dots will ‘leap off the page’ and be detected prior to any conscious, attentive processing on the part of the individual.10
Table 1. US documents that mandate and define prominence and conspicuousness.

<table>
<thead>
<tr>
<th>Document</th>
<th>Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 CFR Section 207.35(b)(3)</td>
<td>Prescription drugs</td>
<td>Requests the presence of an NDC on all drug labels and other drug labelling including the label of any prescription drug container furnished to a consumer. If the NDC is shown on a label, it shall appear ‘prominently’ in limited locations.</td>
</tr>
<tr>
<td>21 CFR Section 201.15 (a)(1–6)</td>
<td>Drugs</td>
<td>General labelling provisions – Reasons for failure with regard to prominence and conspicuousness. Failure to appear. Failure to present in enough places. Failure to extend over the available area. Insufficiency of space for prominent placement (for varied reasons, including devoting greater conspicuousness to other label elements). Size or style of type, insufficient contrast, obscuring designs, crowding.</td>
</tr>
<tr>
<td>Medical Device User Fee and Modernization Act (MDUFMA) of 2002 (Public Law 107-250) – Amends Section 502 of the Federal Food Drug and Cosmetic Act Guidance for Industry and FDA staff; compliance with Section 301 of the MDUFMA of 2002, as amended – Prominent and Conspicuous Mark of Manufacturers on Single-Use Devices.</td>
<td>Device</td>
<td>Requires a device or attachment to the device to bear prominently and conspicuously the name of the manufacturer, a generally recognized abbreviation of such name or a unique and generally recognized symbol identifying the manufacturer.</td>
</tr>
<tr>
<td>27 CFR Chapter 8 Subchapter II Section 215b</td>
<td>Alcoholic beverages</td>
<td>Surgeon General’s required warning ‘GOVERNMENT WARNING: (1) According to the Surgeon General, women should not drink alcoholic beverages during pregnancy because of the risk of birth defects (2) consumption of alcoholic beverages impairs your ability to drive a car or operate machinery, and may cause health problems’ is required to be ‘conspicuous and prominently’ placed on the container in ‘contrasting background’</td>
</tr>
<tr>
<td>27 CFR Section 16.22</td>
<td>Alcoholic beverages</td>
<td>Required warning shall be stated on brand label or separate front label, back or side panel. Separate from other information. Readily legible under ordinary use conditions. Statement beyond ‘GOVERNMENT WARNING’ may not appear in bold. Remaining text shall not be compressed in such a manner that the warning is not readily legible. Containers of 8 fl oz or less = text not less than 1 mm. 8 fl oz–31 = not smaller than 2 mm. Greater than 31 = not smaller than 3 mm.</td>
</tr>
<tr>
<td>40 CFR Section 156.10</td>
<td>Pesticides</td>
<td>All words, statements, graphic representations, designs or other information required on the labelling by the Act or the regulations in this part must be clearly legible to a person with normal vision and must be placed with such conspicuousness (as compared with other words, statements, designs or graphic matter on the labelling) and expressed in such terms as to render it likely to be read and understood by the ordinary individual under customary conditions of purchase and use. Six points or larger type. Clear and contrasting background. Not obscured or crowded.</td>
</tr>
</tbody>
</table>

OBJECTIVES

1. To investigate the conspicuousness of the warning (stage 1 of the information processing model) relative to other label elements, specifically the brand name, strength, net contents and product benefit.
2. To explore how varying graphic elements (an outline and background fill) impact its noticeability when the placement and surface area of the warning are held constant.
3. To develop preliminary data for power analysis that would guide decisions related to sample size in future studies.

MATERIALS AND METHODS

Test participants. Useable data were collected from 17 participants from a population that consisted primarily of students between the ages 18 and 25 years following procedures approved under IRB 06–054. Participants were at least 18 years or older and were not legally blind. Upon arrival, participants were assigned a ‘participant number’ and were characterized by gender and visual acuity.

Visual acuity test. Each participant’s visual acuity was tested and recorded using a near point visual acuity card manufactured by Dow Corning Ophthalmic. The subjects were asked to hold the card approximately 16 in. from their eyes and were asked to read the lowest line possible. This line corresponded with the subject’s near point visual acuity (e.g. 20/20 and 20/30).

Instrument – Applied Science Laboratories eye tracker. An Applied Science Laboratories 501 head mounted bright pupil system (Boston, MA, USA) was used to record participant’s eye movements. Once the eye tracking equipment was comfortably placed on the head, the calibration sequence began. For calibration and testing, subjects were seated at a special table fixtured with a glass pane and a chin rest. This set-up allowed subjects to examine packages at a consistent and fixed distance from their eyes. The calibration sequence employed a floating technique that consisted of nine dots distributed equally across the pane of glass. By fixing the distance and limiting subject movement, the set-up minimized parallax error, enhancing accuracy of the tracking of the gaze trail on the package surface.

Following calibration, participants were instructed that they were being asked to examine a series of items that appeared on a shopping list that they had been given (see Table 2). They were asked to review packages as they would in the store when making a purchase decision for their household (summary of the experimental design shown in Figure 1). They were told that they could view any side of the package that they wished but were instructed to press packages against the glass pane (the calibrated plane) as they did.

Following this instruction, nine packages were handed to the subject for viewing in an order that was counterbalanced across subjects. Participants viewed each of the nine packages for a time period of 10 s before being handed the next product. Ten seconds was chosen as a conservative estimate because sources indicate that consumers generally take 5–7 s examining packages while shopping.11

<table>
<thead>
<tr>
<th>Number</th>
<th>Brand</th>
<th>Package</th>
<th>Size</th>
<th>Warning characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Text colour</td>
</tr>
<tr>
<td>1</td>
<td>Vanish</td>
<td>Toilet bowl cleaner</td>
<td>1.7 oz</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>Tyvedron</td>
<td>Pain reliever</td>
<td>24 gel caps</td>
<td>White</td>
</tr>
<tr>
<td>3</td>
<td>Nature Valley</td>
<td>Almond granola bars</td>
<td>7.4 oz</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>Tyvedron</td>
<td>Pain reliever</td>
<td>24 gel caps</td>
<td>White</td>
</tr>
<tr>
<td>5</td>
<td>Grand River Falls</td>
<td>Tomato basil crackers</td>
<td>1.75 oz</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>Tyvedro</td>
<td>Pain reliever</td>
<td>24 gel caps</td>
<td>Dark blue</td>
</tr>
<tr>
<td>7</td>
<td>Elsa’s Story</td>
<td>Black pepper crackers</td>
<td>5.29 oz</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>Tyvedron</td>
<td>Pain reliever</td>
<td>24 gel caps</td>
<td>Dark blue</td>
</tr>
<tr>
<td>9</td>
<td>PG Tips</td>
<td>Tea bags</td>
<td>4.4 oz</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 2. The nine packages used for the experiment.
This is also consistent with the studies performed by Krugman et al. where subjects were eye tracked while viewing tobacco advertisements. The time spent by the average reader ranged between 3.2 and 14.7 s when they viewed the advertisements without any time restrictions.12

**Stimulus material.** As mentioned, the size, placement and content of the warning on the PDPs were held constant (see Figure 2); the warning design was not. The warning text was comprised of eight-point Swiss 911 1cm BT font in navy blue (Pantone Matching System 184–1). Warnings were created to test for an effect of two design elements, outline and fill, on the ability to garner subject attention. Two levels of each of the elements were crossed such that four designs were tested (See Figure 2 and 3):

- Warning 1 (no outline and no fill).
- Warning 2 (outline and no fill).
- Warning 3 (no outline and fill).
- Warning 4 (outline and fill).

**Statistical analysis.** Eye tracking data were collected in the form of video files and analyzed using Gaze Tracker® (purchased through Applied Science Laboratories, Boston, MA) eye tracking analysis software. All dependent variables tested were first examined for an effect of run order; no effect of order was suggested at $\alpha = 0.05$ for all variables. During the analysis phase of the experiment, five zones (see Figure 4) were created on all four stimulus packages (see Figure 2).
Figure 2. Principal display panels of the four Tyvedron packages used as stimulus.

The five zones were brand name, warning (four levels), net content, strength and product benefit (i.e. nondrowsy). Brand name, net content, strength and product benefit were identical on all four packages of Tyvedron used in the study. Only the warning zone was changed (see Figures 2 and 3). Three dependent variables were used in the analyses:

1. Time spent in zone (a variable response).
2. Probability of a zone being hit first (zone hit first; yes/no; a binary response).
3. The number of visual hits to a zone (a discrete response variable).

Statistical analysis was then carried out using SAS version 9.2 (SAS Institute Inc., Cary, NC, USA). Zone, package, gender and their interactions were included in the model as fixed independent variables, and subject was considered a random factor in the analysis of variance. Time in zone, probability of a zone being hit first and visual ‘hits’ to a zone were considered response variables. Normality and equality of variances assumptions were checked visually using normal plots and side-by-side box plots. Because of right skewness, both time and hits had to be transformed (log-transformed for time and square root for hits). Least square mean estimates and 95% confidence intervals were carried out for the sake of post-hoc comparisons.

RESULTS AND DISCUSSION

Dependent variable – total time spent in a zone

When time was the dependent variable, zone ($p < 0.0001$), zone x gender ($p = 0.0527$) and package x gender ($p = 0.0300$) were found to be significant ($z = 0.05$). Consistent with previous work, brand name garnered significantly more attention than any other aspect of the label when time in zone was
Figure 3. Design of warnings (dimensions are in millimetre; those marked with an asterisk correspond to the centre of the outline).

Figure 4. Zones of test stimulus.

the dependent variable of interest (see Figure 5). The zone warning received significantly less attention than all other zones tested except for strength, only garnering an estimated 0.03 s per viewing.

The analysis of variance also indicated a marginal effect of the interaction term zone by gender on time per zone ($p=0.0527$). To further explore this effect, pairwise comparisons were made using least square mean differences. When all five zones were examined for gender effects, only zone 1, warning, was suggested to be significant ($p=0.0344$). Male subjects spent more time (0.08 s) on the warning zone when compared with their female (0.01 s) counterparts. This contradicts prior research specifically designed to research the effect of gender on the noticeability of warning labels. The reviewed literature suggests female subjects have a greater tendency to look at warnings than male counterparts.14–16
Total time in zone 1 – warning. To specifically explore how varying the graphic and textual elements (an outline and background fill) impacted the time spent on the warning label when its placement and surface area were held constant, we conducted a second analysis that just looked at the time spent on zone 1 (the warnings) (see Table 3) in the form of a $2 \times 2$ factorial that considered fill and outline each at two levels as well as the interaction term. Because of the failure of the data to meet model assumptions, data were log-transformed for the statistical analysis. This analysis did not provide evidence of an effect of warning design on the time spent in the warning zone ($p > 0.05$).

### Table 3. Time spent on Zone 1 – warning (averaged across all the subjects).

<table>
<thead>
<tr>
<th>Warning</th>
<th>Treatment</th>
<th>Total time spent (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warning 1 (no outline, no fill)</td>
<td>Warning 2 (outline, no fill)</td>
<td>4.0</td>
</tr>
<tr>
<td>Warning 3 (no outline, fill)</td>
<td>Warning 4 (outline, fill)</td>
<td>3.9</td>
</tr>
<tr>
<td>Warning 4 (outline, fill)</td>
<td>6.3</td>
<td></td>
</tr>
</tbody>
</table>

Probability of a zone being hit first

We also analyzed the ability of each zone to draw the eye, i.e. the probability of a zone being hit first (see Table 4). Statistical analysis indicated that the probability of being visually ‘hit first’ was significantly affected by the zone ($p < 0.0001$); i.e. some zones were more likely to be hit first than others. No other factors provided evidence of a significant difference ($z = 0.05$). Post-hoc pairwise comparisons suggest that brand name (zone 2) had a significantly higher probability of being hit first (52%) than any other zone ($z = 0.05$) (see Figure 6).
Table 4. Frequency table: number of first hits in a given zone by package.

<table>
<thead>
<tr>
<th>Zone</th>
<th>Package with warning 1 (no outline, no fill)</th>
<th>Package with warning 2 (outline, no fill)</th>
<th>Package with warning 3 (no outline, fill)</th>
<th>Package with warning 4 (outline, fill)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warning (Zone 1)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Brand name (Zone 2)</td>
<td>12</td>
<td>10</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Strength (Zone 3)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Product benefit (Zone 4)</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Net content (Zone 5)</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Please note that eye tracking data were not available for one subject viewing the package with warning 2.

Figure 6. Probability of a zone to be hit first.

Visual hits in a zone

Further analysis examined data for an effect on the number of visual hits garnered. As before, the factors for consideration were subject, zone, package and gender, as well as all interactions. Only subject was considered a random effect. Zone had a significant effect on the number of hits \((p < 0.0001)\), and the interaction term package \(\times\) gender was also indicated to be significant \((p = 0.0043)\). Post-hoc, least square mean comparisons suggested that the brand name (zone 2) received a statistically significant greater amount of attention when the number of hits to the zone was the dependent variable (See

Figure 7. Estimate of number of hits in a zone.

Table 5. Frequency table: number of hits and average number of hits on Zone 1 – warning (across all subjects).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Total number of hits by Warning type (for all subjects)</th>
<th>Average number of hits per subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warning 1 (no outline, no fill)</td>
<td>23</td>
<td>1.35</td>
</tr>
<tr>
<td>Warning 2 (outline, no fill)</td>
<td>21</td>
<td>1.31</td>
</tr>
<tr>
<td>Warning 3 (no outline, fill)</td>
<td>23</td>
<td>1.35</td>
</tr>
<tr>
<td>Warning 4 (outline, fill)</td>
<td>29</td>
<td>1.71</td>
</tr>
</tbody>
</table>
Figure 7). Strength and product benefit were preferentially attended as compared with warnings and net contents at $z = 0.05$.

**Number of hits with respect to the treatments in zone 1 – warning.** Data were also analyzed to explore objective 2, whether or not graphic elements could be manipulated to create a warning that garnered more attention, despite being constrained by placement and area. Preliminary analysis of zone 1 data indicated that when the dependent variable was the probability of first hit, only five subjects (two men and three women) in total hit the warning zone (zone 1) first (see Table 4). Of these, warnings 3 and 4 (no outline and fill and outline and fill, respectively) recorded two hits each, whereas warning 2 (outline and no fill) was hit first once. The frequency with which people hit each of the four warning zones was also examined. This data are presented in Table 5.

Data were too limited to provide meaningful statistical analysis. Although the limited data preclude the ability to provide conclusive results regarding the effect of outline and fill to garner attention, we can observe the following: there were no first hits to warning 1, no outline/no fill (see Figures 2 and, 3 and Table 4). Warnings similar in design to this (no outline and no fill) are currently being used on many of the OTC drugs in the market that are not child resistant.

**CONCLUSIONS**

Our study comprised of three objectives:

1. To test the relative prominence of a warning required by US law to be conspicuous when compared with other elements of the same label.
2. To explore how varying the graphic elements (an outline and background fill) impacted its noticeability when the placement and surface area of the warning were held constant.
3. To develop preliminary data to be used in a power analysis for future studies.

**Objective 1.** Consistent with other work, the warning was less noticeable than brand name for all variables measured (time in zone, probability of first hit and number of hits to zone) at $z = 0.05$. Comparisons of the warning with the other zones (strength, net contents and product benefits) were more varied (see Figures 5–7).

**Objective 2.** With regard to inferences relating to objective 2, the ability of varied graphic elements to garner consumer attention, our sample size ($n = 17$) severely limited our ability to make statistical inferences. This was despite the fact that it was typical for an eye tracking study. Regardless, the fact that the warnings designs were trending in the expected direction (see Tables 4 and 5) encourages future study. Before future study can begin, a power analysis is needed.

**Objective 3.** Power analyses (see Figure 8) were performed. Between subject, variance ($\sigma^2 = 38.69$) was considered equal; two expected mean differences (1.5 and 2.5 s) for the time spent in zone 1 (warning) were used in the power analysis. The analysis indicates that a sample size of over 980
subjects would be required for expected mean difference of 1.5 s and that 190 subjects would be required for 2.5 s to achieve 80% power.

That said, this study provides several interesting implications for future research. Eye tracking as a method for the objective evaluation of attention has not been used very often in the field of packaging, and of these studies, several track eye movement as subjects view a label on a computer screen, as opposed to allowing them to freely manipulate the package. Tracking on a computer screen requires a flattened image of the graphic information be presented, forcing exposure to all six faces of the package. Our techniques allow subjects to freely manipulate the package by rotating it to any face that they desire. As such, the study presented herein is among the first to track subjects on real packages, providing a more ecologically valid context for the collection of data regarding attentive behaviour.

This represents a methodology that has the potential to accurately characterize the attentive behaviours of individuals, thereby objectively evaluating the effectiveness of varied elements of package design.

LIMITATIONS

We purposefully limited our study’s scope such that it examined the ability of a warning, required by law to be conspicuously displayed, to garner attention as measured with our Applied Science Laboratories 501 eye tracking device. Although attention is requisite to comprehension and action, it is but a single step in information processing. As such, we are limited in our ability to conjecture about the success or failure of the message in conveying information. Additionally, it should be noted that other factors, not collected or reported herein, have the potential to significantly influence the attentive behaviours of subjects. For instance, Sansgiry and Cady found that elderly consumers ‘were more involved in the decision making process to purchase OTC medications . . . The elderly not only purchase and spend more money on medications but also read OTC labels completely.’ Key factors, such as drug familiarity and consumer involvement with the drug, were not recorded and could have a significant effect in findings. Additionally, our subject population, composed of college students, is not generalizable.

ACKNOWLEDGEMENTS

We are grateful for the contribution of student workers; Scott Merriman, who collected much of the data presented, and Amanda Wyman, who was actively involved with some of the data analysis. We gratefully acknowledge the help received from Kateryna Ananyeva, statistical consultant at the Statistical Consulting Center (SSC) in the College of Agriculture and Natural Resources at Michigan State University, for supervision and support of the analysis of the data. We would also like to acknowledge the editorial contributions of Dr Susan Selke.

REFERENCES