Background: The Family Smoking Prevention and Control Act gave the U.S. Food and Drug Administration (FDA) legal authority to mandate graphic warning labels on cigarette advertising and packaging. The FDA requires that these graphic warning labels be embedded into cigarette advertising and packaging by September 2012.

Purpose: The aim of this study was to examine differences in recall and viewing patterns of text-only versus graphic cigarette warning labels and the association between viewing patterns and recall.

Methods: Participants (current daily smokers; N = 200) were randomized to view a cigarette advertisement with either text-only or graphic warning labels. Viewing patterns were measured using eye-tracking, and recall was later assessed. Sessions were conducted between November 2008 and November 2009. Data analysis was conducted between March 2011 and July 2011.

Results: There was a significant difference in percentage correct recall of the warning label between those in the text-only versus graphic warning label condition, 50% vs 83% ($\chi^2=23.74, p=0.0001$). Time to first viewing of the graphic warning label text and dwell time duration (i.e., time spent looking) on the graphic image were significantly associated with correct recall. Warning labels that drew attention more quickly and resulted in longer dwell times were associated with better recall.

Conclusions: Graphic warning labels improve smokers' recall of warning and health risks; these labels do so by drawing and holding attention.

In November 2010, the FDA released a collection of 72 advertisement versions (36 images), from which nine final versions were identified in June 2011, required by law to appear on every cigarette advertisement and package by September 2012. Litigation filed in August 2011 by four tobacco companies claims that these warning labels violate their First Amendment rights and will be costly. The companies subsequently requested that the introduction of the warning labels be stopped. The warning labels were temporarily blocked in November 2011, and litigation is likely to continue.

Graphic warning labels have proven to be effective in eliciting negative responses to smoking in U.S. smokers, increasing reported intention to quit smoking in Canada, and modifying beliefs about smoking dangers. However, research on cigarette warning labels has generally been conducted using large, population-based studies that could be confounded by concurrent tax increases or anti-smoking media campaigns that coincide with the
introduction of new warning labels. Also, measures of attitudes and beliefs can be confounded by participant expectancy (e.g., participants may rate the novel warning more favorably, or report greater intention to quit smoking, because they deduce the study aim).

An important first step in evaluating whether a warning label is effective is to demonstrate whether smokers can correctly recall its content or message. Previous research using eye-tracking to evaluate the effectiveness of cigarette warning labels has been limited, but one study reported that adolescents’ viewing time of text-only warning labels was associated with a masked recall task (i.e., identifying the blacked-out area as a warning label post-advertisement viewing). But such viewing time was not associated with free recall, such as stating the warning label content, an objective measure that often has been used and validated in health communications research. Recently, Munafò et al. reported greater visual attention to graphic warning labels with plain packaging compared to cigarette brand packaging, but only in nondaily smokers. Attention and memory theory suggests that viewing duration is positively associated with recall, but as Munafò et al. illustrate, attention to the rest of a cigarette package may distract viewers from a warning label.

One advantage of understanding the relationship between viewing patterns and ability to recall is that it may elucidate specific aspects of the warning label that make it effective. Results can then empirically inform future warning label development, in keeping with provisions of the FSPCA to rotate and improve warning labels used in the U.S. The purpose of the current study is to evaluate the effect of embedding a graphic warning label into a popular cigarette advertisement on ability to recall the warning label content, and to examine if and how viewing patterns affect recall ability. Recall was an open-ended free-recall item. It was hypothesized that recall would be greater for the graphic warning label advertisement than for the text-only warning label, that measures of viewing patterns would be greater in the graphic warning labels compared to the text-only warning labels, and that increased viewing of the graphic warning label would be associated with correct recall.

Methods

Participants

Participants were 200 current smokers who self-reported smoking a minimum of ten cigarettes daily for a minimum of 5 years; were not currently trying to quit smoking; were aged 21–65 years; spoke English as a primary language; were not current substance abusers; had no visual impairments including wearing glasses (contact lenses acceptable) or color blindness; and who were willing to attend a laboratory session to complete questionnaires related to smoking history and tobacco products. Participants were a convenience sample who responded to newsprint advertisements and local online recruitment postings in the greater Philadelphia area, and were not intended to be representative of the smoking population. Participants received $50.00 in remuneration for completing the 75-minute session. The protocol was approved by the University of Pennsylvania IRB; sessions were conducted during November 2008 – November 2009, prior to the FDA’s release of the testing messages, and analyses were conducted March 2011–July 2011.

Procedures

All sessions were conducted in a media research laboratory room within a nicotine addiction research clinic at a university medical school center. Participants had study procedures explained, provided signed informed consent, smoked a cigarette to standardize time since last cigarette, completed demographic and smoking history measures, and then had the eye-tracking equipment calibrated while sitting in a straight-back chair 1.0 m from a 19-inch computer monitor.

Participants were randomized to view either the text-only (Text condition) warning label advertisement, which was unaltered and utilized the Surgeon General’s warning and Federal Trade Commission (FTC) testing information that has appeared on cigarette advertisements since 1985; or the graphic (Graphic condition) warning label version that was one version of graphic warning employed by Health Canada and scaled to occupy 22% of the cigarette advertisement space. Figure 1 shows a modified version of the Text condition and Graphic condition images used in this study. This percentage of coverage was consistent with the range other countries were adopting at the time this protocol was developed and predicted image distortion. Participants were instructed to view the advertisement for the entire presentation time, 30 seconds to have ample time to view and process materials. Following viewing the advertisement, participants completed two distractor items about their general opinions of the advertisement (What thoughts [opinions] did you have about the image?), followed by the warning label recall question.

Measures

Demographic and smoking history measures. Age, gender, race, ethnicity, education level, years smoking, preferred cigarette brand, daily cigarette consumption, and nicotine dependence was assessed by self-report. Race and ethnicity were assessed using standard NIH guidelines. Nicotine dependence was assessed using the Fagerstrom Test of Nicotine Dependence (FTND), a six-item self-report measure of nicotine dependence (range 0–10) with satisfactory internal consistency and high test–retest reliability.

Outcome

Recall. Recall was assessed by asking participants, Based on the ad you just viewed, please type what the warning label read. Three trained raters unaware of the study hypothesis scored each participant statement as correct or incorrect. Correct was operationally defined as containing the words: quitting, smoking, reduces, risk, and health, or the root words (e.g., reduce), for the Text condition warning label; and, warning, lung, cancer, smoke, increases, or the root words, for the Graphic condition. Incorrect was defined as any response that contained fewer than five of the target words in each set. In the event of scoring inconsistencies, the project principal investigator made a final scoring determination (n = 2, coded incorrect). This recall measure and its scoring is derived from standard open-ended recall procedures and is similar in format and
duration after presentation to one of the recall items used in assessing the efficacy of the proposed warning labels.²³

Eye-tracking. Gazetracker software (v.07.01.243.128) was used to display the advertisements and record the eye-tracking data as measured by an Eye-Trac 6 control unit with an R6 pan/tilt optics system and video head tracker (Applied Science Laboratories, Boston, MA). Data transfer was linked using the Eye-Trac 6 User Interface Program (v.1.30.8.0).

Areas of interest were identified a priori in each ad, and consisted of whole ad (to verify viewing screen); product name; cowboy image; and warning area. The Surgeon General’s warning box and FTC text were coded as areas of interest for the Text condition (Figure 1b), and the graphic image and warning text for the Graphic condition (Figure 1a). For each region, measures of interest were dwell time (total time viewed in area of interest; seconds); fixations (number of times each area of interest was observed; frequency); and time to first viewing of area of interest (seconds), as these objective measures are associated with attention, reading, and information processing.¹²,¹⁶,²⁴ Fixations were operationalized as any 60-pixel-diameter space with three consecutively sampled observations with a minimum cumulative 200-ms duration, consistent with previous viewing images and reading tasks.²⁴

Analysis Plan
Study sample was characterized using descriptive statistics and t-tests or chi-square used to identify potential group differences. Participant responses to the open-ended recall item were scored as correct or incorrect and evaluated for reliability with Krippendorff’s alpha.²⁵ Chi-square was used to examine differences in recall between conditions. Participant eye-tracking data underwent data reduction and processing, which subsequently produced measures of interest for each area of interest. ANCOVA was used to examine condition differences in eye-tracking data. Logistic regression analysis was conducted to determine models of predicting correct recall of warning label content for Graphic and Text condition advertisements. Descriptive and eye-tracking measures were inserted into each model, and those with p value >0.2 were subsequently excluded.

Results
Demographic and Smoking History Measures
The participant sample (n=200) was 64.5% male, and 92% Caucasian, 3.6% Asian, 3.1% African-American, and 1.3% more than one race; three individuals reported Hispanic ethnicity irrespective of race. Most participants had completed a high school degree (98%); 40% had completed college. The participant sample was an average age of 30.45 (SD=8.0) years; reported smoking for 12.8 years (SD=6.8); and smoked an average of 16.6 cigarettes daily (SD=5.8), with an average FTND nicotine dependence score of 5.1 (SD=2.1). The most popular brands were Marlboro (41%); Camel (25%); Parliament (15%); and American Spirit (8%). The
reported number of daily cigarettes was the only descriptive measure to differ significantly between Text and Graphic conditions: 17.7 (SD=6.1) vs 15.5 (SD=5.2), p=0.008, respectively, and therefore was retained as a covariate in subsequent analyses.

Recall
There was a significant difference in the percentage of participants with correct recall of the warning label between Text and Graphic conditions: 50% vs 83% (χ²=23.74, p=0.0001), respectively. Krippendorff’s alpha for interrater reliability for scoring the recall item was excellent (α=0.914).

Eye-Tracking
Eye-tracking results are presented in Table 1. There were eight sessions with low-quality eye-tracking data (five Text, three Graphic) that were excluded from analysis of eye-tracking data. Viewing patterns of the whole advertisement were examined to verify attention to the screen during presentation, and were found to not differ by condition. The warning area in the Graphic condition was viewed for a significantly greater duration than in the Text condition (11.1 vs 8.2, p<0.001), and time to first viewing was significantly shorter (1.7 vs 2.5, p<0.002). There was a trend toward the cowboy image being viewed for a shorter duration and with fewer fixations in the Graphic condition compared to the Text condition (Table 1).

### Table 1. Eye-tracking measures for cigarette advertisements with Text condition or Graphic condition warning labels

<table>
<thead>
<tr>
<th>Area of interest and measures</th>
<th>Text condition (n=95)</th>
<th>Graphic condition (n=97)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole ad (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>26.3 (5.2)</td>
<td>25.2 (5.6)</td>
<td>0.16</td>
</tr>
<tr>
<td>Fixations</td>
<td>19.5 (20.6)</td>
<td>16.6 (19.3)</td>
<td>0.32</td>
</tr>
<tr>
<td>Time to first view</td>
<td>0.04 (0.09)</td>
<td>0.06 (0.16)</td>
<td>0.36</td>
</tr>
<tr>
<td>Warning area</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>8.2 (5.4)</td>
<td>11.1 (4.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fixations</td>
<td>10.8 (13.1)</td>
<td>9.6 (12.3)</td>
<td>0.51</td>
</tr>
<tr>
<td>Time to first view</td>
<td>2.5 (2.1)</td>
<td>1.7 (1.2)</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td>Brand name (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>1.6 (1.3)</td>
<td>1.7 (1.3)</td>
<td>0.63</td>
</tr>
<tr>
<td>Fixations</td>
<td>1.4 (2.5)</td>
<td>1.4 (2.6)</td>
<td>0.97</td>
</tr>
<tr>
<td>Time to first view</td>
<td>1.5 (4.3)</td>
<td>0.8 (1.4)</td>
<td>0.15</td>
</tr>
<tr>
<td>Cowboy image (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>8.8 (4.8)</td>
<td>7.7 (3.6)</td>
<td>0.07</td>
</tr>
<tr>
<td>Fixations</td>
<td>6.5 (8.9)</td>
<td>4.4 (6.6)</td>
<td>0.06</td>
</tr>
<tr>
<td>Time to first view</td>
<td>0.1 (.13)</td>
<td>0.3 (1.3)</td>
<td>0.15</td>
</tr>
<tr>
<td>Graphic image (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>—</td>
<td>4.2 (2.7)</td>
<td>—</td>
</tr>
<tr>
<td>Fixations</td>
<td>—</td>
<td>4.9 (6.7)</td>
<td>—</td>
</tr>
<tr>
<td>Time to first view</td>
<td>—</td>
<td>4.5 (3.1)</td>
<td>—</td>
</tr>
<tr>
<td>Graphic text (5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>—</td>
<td>6.9 (3.6)</td>
<td>—</td>
</tr>
<tr>
<td>Fixations</td>
<td>—</td>
<td>2.6 (4.2)</td>
<td>—</td>
</tr>
<tr>
<td>Time to first view</td>
<td>—</td>
<td>1.7 (1.3)</td>
<td>—</td>
</tr>
<tr>
<td>Surgeon General’s warning (6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>2.8 (2.2)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Fixations</td>
<td>2.6 (4.2)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Time to first view</td>
<td>3.4 (2.6)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>FTC testing statement (7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwell time</td>
<td>5.3 (4.1)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Fixations</td>
<td>6.5 (8.6)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Time to first view</td>
<td>3.1 (2.5)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note: Values are M (SD), with significance values where appropriate. Numbers in parentheses in area of interest correspond to numbers in Figure 1 delineating sections. The warning area is defined as the merging of the areas labeled 6 and 7 for the Text condition image, and 4 and 5 for the Graphic condition image. FTC, Federal Trade Commission
Association Between Eye-Tracking and Recall

In the logistic regression model for Graphic with recall as the dependent measure, time to first viewing of the graphic warning label text (OR = 0.41, 95% CI = 0.22, 0.76); dwell time for the graphic image (OR = 1.37, 95% CI = 1.00, 1.86); and preferred cigarette brand (OR = 0.18, 95% CI = 0.04, 0.81) were significantly associated with correct recall, in a model that also included daily cigarette consumption (OR = 1.11, 95% CI = 0.96, 1.3, p = 0.17). Shorter time for attention to be drawn to the warning label text, longer dwell time on the graphic image, and not being a Marlboro smoker (the cigarette brand in the advertisement) were associated with correct recall.

There were no significant predictors of recall in the logistic regression model for Text. However, dwell time of whole screen (OR = 1.06, 95% CI = 0.99, 1.14, p = 0.09) and number of fixations of brand name (OR = 0.86, 95% CI = 0.71, 1.02, p = 0.10) had trends toward significant associations with correct recall. Greater dwell time on the entire screen and fewer fixations on the brand name were not associated with correct recall.

Discussion

The current study is the first to specifically evaluate how adult U.S. smokers view graphic warning labels in cigarette advertising, in conjunction with how these viewing patterns affect ability to recall the warning label content. Recall was significantly greater for the graphic warning labels compared to the text-only labels, which is consistent with findings of increased knowledge of cigarette smoking risks in countries where graphic warning labels are employed. As noted in Hammond et al., knowledge can arise from several sources, and free or unprompted recall likely has a higher threshold for assessing risk perception than other common measures. Therefore, the authors interpret the high (83%) recall of warning label text in the graphic condition with enthusiasm. It is relatively challenging to correctly respond to this recall item, yet smokers were very capable of doing so in the Graphic as opposed to the Text condition (50%). This clearly demonstrates the effectiveness of, and support for, implementing graphic warning labels into cigarette advertisements and packaging in the U.S.

The present study also provides valuable insight into how the warning labels may be effective, which may serve to create more effective warning labels in the future. The Text condition warning label was not recalled correctly as frequently as that in the Graphic condition, and viewing patterns suggest that viewing duration across the whole advertisement and fewer fixations on the largest text (i.e., the brand name) were weakly associated with recall. Comprehension of reading text alone is often associated with viewing time and fixations and is consistent with the current results. In contrast, the Graphic condition warning label area quickly drew attention and held attention, relative to that in the Text condition, and these measures were associated with correct recall.

One potential explanation for the findings is the possibility that smokers were already familiar with the text in the Text condition and therefore spent less time inspecting it. However, this would not explain why they could not recall the text as well as those in the Graphic condition. Further, total time spent viewing the warning in the Text condition did not correlate with recall, whereas it did in the Graphic condition. This finding is consistent with research indicating that pictorial information that serves to complement textual information enhances memory for the text. This is an important feature of graphic warnings that can make them more effective than the text-only warnings that are currently employed in the U.S.

Results from the current study provide important policy information regarding the development of future warning labels. Time to first viewing is associated with attention, and in the case of a cigarette advertisement, attracting attention to a health warning may be important because it distracts from viewing other parts of the advertisement. Prior research has shown that the body of a cigarette advertisement can alter smokers’ risk perceptions of a cigarette product using explicit and implicit manipulations.

The current study demonstrates that drawing attention to the warning label can improve recall of health-relevant information, which may extend to improving risk perception of smoking. Additionally, attracting attention to the warning before viewing the advertisement body may change the framing of the message in the advertisement body, causing viewers to approach it with greater caution. Finally, time to first viewing has good application to real-world settings where people may allocate only a few seconds to a print advertisement. Further study on the size, font, color, and location of text may identify the most effective way to draw attention.

Considerable research exists on how smokers rate graphic labels for acceptance, vividness, and ability to evoke strong emotional responses. In general, U.S. smokers are accepting of the images, which have considerably less gore than graphic warning labels used in other countries. As noted by Dr. Lawrence Dayton, MD, Director of the FDA’s Center for Tobacco Products, “sometimes the images that work best are the ones that people can look at and have an emotional impact but not dismiss.” The present study empirically supports that claim: As viewing duration of the graphic image increases, the likelihood of correct recall increases. The use...
of this specific graphic image in the warning label increased the ability of smokers to recall the health message because of their viewing duration.

An unexpected finding was that being a Marlboro smoker was significantly associated with being less likely to correctly recall the Graphic condition warning label (75% vs 88%). This could potentially be attributable to Marlboro smokers viewing images relevant to their preferred brand more, or the warning area less, than smokers of other brands. The former possibility supports banning advertisement content, as some have proposed, including plain or blank packaging, whereas the latter suggests that regulating the body of the advertisement to include health-related information might be an effective strategy because this area is being viewed more attentively. The study was not designed or powered to properly test this empirical question, but it certainly deserves consideration and could inform regulatory stakeholders. Further research also should consider how brand-switching affects risk knowledge and perception, as well as how those contemplating trying cigarette smoking view and understand cigarette advertisements and warning labels.

The current study does have a few limitations. Recall occurred within only a few minutes after viewing the advertisement, and future work on graphic warning labels should include the effect that chronic exposure to graphic warning labels has on long-term recall, changes in attitude toward smoking, and changes in behavior, such as increased smoking-cessation attempts, similar to effects reported by Hammond et al., after the introduction of graphic warning labels. However, Hammond et al. did observe high correct recall and depth of processing, and the latter was associated with intention to quit, supporting the usefulness of short-term recall. Further, short-term recall was assessed in a recent experimental study of graphic cigarette warning labels, in addition to recall after 1 week and cognitive and emotional reactions, and results suggest a similar trend in relation to the recall items. As noted by Nunnemaker et al., immediate responses (i.e., emotional and recall) likely influence short-term responses (i.e., 1-week recall, attitudes) which eventually lead to changes in quitting intentions and not starting to smoke.

Also, the warning areas were different between conditions: 123,246 vs 62,622 pixels, 16 vs 13 words, and greater text size for the Graphic and Text conditions, respectively. Therefore, group comparisons must be interpreted with caution. However, the warning labels were sized to currently mandated format for the Text condition, and pending required size for the Graphic condition; thus, results convey important policy-relevant information. Clearly, evaluation of several types is necessary to elucidate themes or styles that are most effective, as has been done with anti-smoking public service announcements.

Results from the current study demonstrate the effectiveness of graphic warning labels in cigarette advertisements in increasing recall of warning label-based smoking risks and provide novel objective evidence that smokers’ viewing patterns of cigarette advertisements containing graphic warning labels are associated with recall. Graphic warning labels should be incorporated into cigarette advertisements without delay; not doing so only prolongs an overdue, necessary improvement to U.S. tobacco control.

This work was supported by Robert Wood Johnson Foundation (62622; AAS), National Cancer Institute R01-120594 (AAS), the National Cancer Institute Center of Excellence in Cancer Communication Research (CECCR) located at the Annenberg School for Communication (P50-CA095856-05 and P20-CA095856-06; RC Hornik, principal investigator), and the NIH (P50-CA143187; CE Lerman, principal investigator).

No financial disclosures were reported by the authors of this paper.

References


Supplementary data
A pubcast created by the authors of this paper can be viewed at www.ajpmonline.org/content/video_pubcasts_collection.

Did you know?
Two AJPM articles per issue offer CME credits.
Go to www.ajpmonline.org/cme/home for more information.