Recall of health warnings in smokeless tobacco ads

L Truitt, W L Hamilton, P R Johnston, C P Bacani, S O Crawford, L Hozik, C Celebucki

Objective: To determine the effects of health warning characteristics in smokeless tobacco magazine print ads on warning recall, and the implications for current US Federal regulations.

Design: Subjects examined two distracter ads and one of nine randomly assigned smokeless tobacco ads varying in health warning presence, size (8 to 18 point font) and contrast (low versus high)—including no health warning. They were then interviewed about ad content using recall and recognition questions.

Subjects: A convenience sample of 895 English speaking males aged 16–24 years old who were intercepted at seven shopping malls throughout Massachusetts during May 2000.

Main outcome measures: Proven aided recall, or recall of a health warning and correct recognition of the warning message among distracters, and false recall.

Results: Controlling for covariates such as education, employment/student status, and Hispanic background, proven aided recall increased significantly with font size; doubling size from 10 to 20 point font would increase recall from 63% to 76%. Although not statistically significant, recall was somewhat better for high contrast warnings. Ten per cent of the sample mistakenly recalled the warning where none existed.

Conclusions: As demonstrated by substantially greater recall among ads that included health warnings over ads that had none, health warnings retained their value to consumers despite years of exposure (that can produce false recall). Larger health warnings would enhance recall, and the proposed model can be used to estimate potential recall that affects communication, perceived health risk, and behavior modification.

Smokeless tobacco includes various ground or cut tobacco products, such as moist oral snuff, dry oral and nasal snuff, or chewing tobacco. According to the US Surgeon General, consequences of smokeless tobacco use include nicotine addiction, oral cancer, gum recession, elevated blood pressure, and cardiovascular disease. Over 41 million Americans (about 19% of the US population) have used smokeless tobacco in their lifetimes; prevalence is highest among persons aged 18–25 years (24%). Non-Hispanic white males in this age group rank first in smokeless tobacco use (16%) over the previous 12 months.

The Comprehensive Smokeless Tobacco Health Education Act of 1986 directs the US Federal Trade Commission to regulate statutory health warnings on smokeless tobacco packaging and advertising. Manufacturers, packagers, and importers must display rotating health warning labels on packages and advertising regarding cancer, gum disease, or cigarette substitution. "WARNING" must appear in a wide arrow that points to the uppercase warning message inside a circle, and may be printed in colors that contrast with the ad background or in larger and bolder type regardless of background.

In addition to demographic, sociopsychological, and other individual differences, message communication has implications for perceived health risk and behavior modification. Most tobacco related studies concern cigarette advertising research, which indicates the impact of Surgeon General health warnings has diminished with overexposure, but this difference failed tests of statistical significance; furthermore, false positives among a control group shown no warning suggests a baseline of dubious responses among treatment groups.

To evaluate the effectiveness of smokeless tobacco health warning requirements, we tested young males—associated with the highest incidence of smokeless tobacco use—randomly assigned to magazine ads that varied in warning presence, contrast, and font size. We tested three hypotheses.

• Does warning size affect recall? Given ad design constraints, do size increases over the currently mandated 10 point font generate significant increases in recall?

• Does the degree of contrast (low versus high) between the warning and ad background significantly affect recall?

• Is recall produced by the mandated 10 point font significantly different from false recall produced by an ad with no warning?

This extends Popper and Murray's research concerning recall to a population of young males who vary in education, employment, and student status, and uses interviewers rather than self administration. The results are model based, rather than discrete measures associated with limited warning manipulations, and permit estimation to a range of warning sizes.

METHODS

Subjects

During May 2000, a convenience sample of 895 males between 16–24 years old (median 19 years) who resided or attended school in Massachusetts were intercepted at seven shopping malls statewide. Of the 27.8% under 18 years old, 96.3% were in high school or trade school. Among those 18 years or older, 38.7% were enrolled in college, and 55.8% were employed full time. Over two thirds (69.6%) were non-Hispanic whites.
Table 1 Proportion of observed aided recall by health warning contrast and font size

<table>
<thead>
<tr>
<th>Contrast</th>
<th>8 point</th>
<th>10 point</th>
<th>14 point</th>
<th>18 point</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample size</td>
<td>Proportion</td>
<td>Sample size</td>
<td>Proportion</td>
</tr>
<tr>
<td>No warning</td>
<td>99</td>
<td>0.10</td>
<td>99</td>
<td>0.10</td>
</tr>
<tr>
<td>Low contrast</td>
<td>-</td>
<td>-</td>
<td>100</td>
<td>0.44</td>
</tr>
<tr>
<td>High contrast</td>
<td>-</td>
<td>-</td>
<td>99</td>
<td>0.59</td>
</tr>
</tbody>
</table>

13.9% were African American, 11.0% were Hispanics, and the rest were Middle Eastern (2.8%), Asian (1.7%), or of another background (1.1%). Most (93.5%) had never married. Nearly a third (31.6%) had noticed a magazine ad for smokeless tobacco within the past 30 days. Current tobacco use rates were 36.4% for cigarettes and 3.8% for smokeless tobacco (compared to 7.4% of males in Massachusetts high schools).^{13}

Print ads
Three professionally prepared colour prints of actual magazine ads were presented on 8.5 by 11 inch (20 cm by 29 cm) paper. The two distracter colour prints used to conceal the focus of the study included one automobile ad with minimal text (no health warnings) against a dark background, and one protein drink ad with detailed text concerning product use (a health warning) against a light background. The Skoal brand smokeless tobacco ad contained minimal text against the sunset image of a man at the beach (see appendix). The mandated less tobacco ad contained minimal text against the sunset warning) against a light background. The Skoal brand smokeless tobacco ad with detailed text concerning product use (a health warning). That there was no health warning was a response (two other mandatory warnings, and a nicotine addiction rect message regarding gum disease among three distracters). Also, subjects who freely recalled the warning symbol or message (Q1) were not asked if a warning message was present (Q2); if the correct message was identified (Q3) and proven unaided recall was demonstrated, this counted toward proven aided recall.

Procedure
Mail intercept service providers recruited 16–24 year old males who resided or attended school in Massachusetts. Interviewers drew materials from packets presorted in repeating series of treatment groups 1 through 9. The interviewer obtained verbal consent and gave the subject three ads to examine for 60 seconds. Interviews lasted 5–10 minutes, ending with the interviewer providing an explanation of the study along with materials about smokeless tobacco and its effects.

RESULTS
Table 1 shows the proportion of the sample who recalled the warning under various contrast and font size combinations. Although these sample proportions are legitimate estimates of their population counterparts, we preferred to obtain our results via a regression model. Model based estimates adjust for covariates, fit smoothly through the random fluctuations of the observed proportions, tend to be more precise, and provide interpolated and extrapolated estimates at font sizes not actually observed.

We used a combination of stepwise methods and all subset model methods to identify additional factors from a list of 31 candidates comprising 10 covariates, 10 covariate-by-font interactions, 10 covariate-by-contrast interactions, and a contrast-by-logfont interaction term. The 10 covariates included: continuous measures of education and age (years); and bivariate measures of Hispanic ethnicity, employment/student status, marital status, cigarette use (past month, smoked 100 in lifetime), smokeless tobacco use (past month), and smokeless tobacco ad exposure. None of the interactions were found to be significant. The final model expressed the probability of recall as a logistic function of contrast, font size, exposure, tobacco use, and demographics.

Recall entails information retrieval that is initiated without explicit external cues, whereas recognition occurs when subjects identify information with external cues (in lists among distracters). Free or unaided recall tasks are generally more difficult than recognition tasks; however, with prompted or aided recall, subjects have a 50% chance of guessing that a warning was present regardless of whether their recall was true. We constructed the more stringent measure of proven aided recall by combining recall with message recognition.* Proven aided recall occurred when the subject recalled a health warning with prompting (Q2) and correctly identified the message among distracters (Q3). Also, subjects who freely recalled the warning symbol or message (Q1) were not asked if a warning message was present (Q2); if the correct message was identified (Q3) and proven unaided recall was demonstrated, this counted toward proven aided recall.

RESULTS
Table 1 shows the proportion of the sample who recalled the warning under various contrast and font size combinations. Although these sample proportions are legitimate estimates of their population counterparts, we preferred to obtain our results via a regression model. Model based estimates adjust for covariates, fit smoothly through the random fluctuations of the observed proportions, tend to be more precise, and provide interpolated and extrapolated estimates at font sizes not actually observed.

We used a combination of stepwise methods and all subset model methods to identify additional factors from a list of 31 candidates comprising 10 covariates, 10 covariate-by-font interactions, 10 covariate-by-contrast interactions, and a contrast-by-logfont interaction term. The 10 covariates included: continuous measures of education and age (years); and bivariate measures of Hispanic ethnicity, employment/student status, marital status, cigarette use (past month, smoked 100 in lifetime), smokeless tobacco use (past month), and smokeless tobacco ad exposure. None of the interactions were found to be significant. The final model expressed the probability of recall as a logistic function of contrast, font size, exposure, tobacco use, and demographics.

Recall entails information retrieval that is initiated without explicit external cues, whereas recognition occurs when subjects identify information with external cues (in lists among distracters). Free or unaided recall tasks are generally more difficult than recognition tasks; however, with prompted or aided recall, subjects have a 50% chance of guessing that a warning was present regardless of whether their recall was true. We constructed the more stringent measure of proven aided recall by combining recall with message recognition.* Proven aided recall occurred when the subject recalled a health warning with prompting (Q2) and correctly identified the message among distracters (Q3). Also, subjects who freely recalled the warning symbol or message (Q1) were not asked if a warning message was present (Q2); if the correct message was identified (Q3) and proven unaided recall was demonstrated, this counted toward proven aided recall.

RESULTS
Table 1 shows the proportion of the sample who recalled the warning under various contrast and font size combinations. Although these sample proportions are legitimate estimates of their population counterparts, we preferred to obtain our results via a regression model. Model based estimates adjust for covariates, fit smoothly through the random fluctuations of the observed proportions, tend to be more precise, and provide interpolated and extrapolated estimates at font sizes not actually observed.

We used a combination of stepwise methods and all subset model methods to identify additional factors from a list of 31 candidates comprising 10 covariates, 10 covariate-by-font interactions, 10 covariate-by-contrast interactions, and a contrast-by-logfont interaction term. The 10 covariates included: continuous measures of education and age (years); and bivariate measures of Hispanic ethnicity, employment/student status, marital status, cigarette use (past month, smoked 100 in lifetime), smokeless tobacco use (past month), and smokeless tobacco ad exposure. None of the interactions were found to be significant. The final model expressed the probability of recall as a logistic function of contrast, font size, exposure, tobacco use, and demographics.

Recall entails information retrieval that is initiated without explicit external cues, whereas recognition occurs when subjects identify information with external cues (in lists among distracters). Free or unaided recall tasks are generally more difficult than recognition tasks; however, with prompted or aided recall, subjects have a 50% chance of guessing that a warning was present regardless of whether their recall was true. We constructed the more stringent measure of proven aided recall by combining recall with message recognition.* Proven aided recall occurred when the subject recalled a health warning with prompting (Q2) and correctly identified the message among distracters (Q3). Also, subjects who freely recalled the warning symbol or message (Q1) were not asked if a warning message was present (Q2); if the correct message was identified (Q3) and proven unaided recall was demonstrated, this counted toward proven aided recall.

RESULTS
Table 1 shows the proportion of the sample who recalled the warning under various contrast and font size combinations. Although these sample proportions are legitimate estimates of their population counterparts, we preferred to obtain our results via a regression model. Model based estimates adjust for covariates, fit smoothly through the random fluctuations of the observed proportions, tend to be more precise, and provide interpolated and extrapolated estimates at font sizes not actually observed.

We used a combination of stepwise methods and all subset model methods to identify additional factors from a list of 31 candidates comprising 10 covariates, 10 covariate-by-font interactions, 10 covariate-by-contrast interactions, and a contrast-by-logfont interaction term. The 10 covariates included: continuous measures of education and age (years); and bivariate measures of Hispanic ethnicity, employment/student status, marital status, cigarette use (past month, smoked 100 in lifetime), smokeless tobacco use (past month), and smokeless tobacco ad exposure. None of the interactions were found to be significant. The final model expressed the probability of recall as a logistic function of contrast, font size, exposure, tobacco use, and demographics.

Recall entails information retrieval that is initiated without explicit external cues, whereas recognition occurs when subjects identify information with external cues (in lists among distracters). Free or unaided recall tasks are generally more difficult than recognition tasks; however, with prompted or aided recall, subjects have a 50% chance of guessing that a warning was present regardless of whether their recall was true. We constructed the more stringent measure of proven aided recall by combining recall with message recognition.* Proven aided recall occurred when the subject recalled a health warning with prompting (Q2) and correctly identified the message among distracters (Q3). Also, subjects who freely recalled the warning symbol or message (Q1) were not asked if a warning message was present (Q2); if the correct message was identified (Q3) and proven unaided recall was demonstrated, this counted toward proven aided recall.
The odds of recall for an employed person (working or studying) was over three times the odds for an unemployed person (p = 0.004) and a one unit (roughly one year) increase in education increased the odds of recall by 10% (p = 0.037). The odds of recall for an Hispanic were 46% less than the odds for a non-Hispanic (p = 0.008).

The effects of contrast and font on the probability (rather than the odds) of recall are shown in table 3 and fig 1. These two curves are the curves over the two, four dimensional surfaces corresponding to the sample average for education (11.95), work (0.96), and Hispanic (0.11). For example, the probability of recall for high contrast warnings increased by 21% from 0.63 to 0.76 as font size increased from 10 to 20.

Averaging over contrasts, the predicted probability of recall at font sizes 8, 10, 14, and 18 was 0.56, 0.61, 0.67, and 0.72, respectively. Predicted probabilities at other font sizes can be calculated from the parameter estimates (table 2), as shown in the upper curve of fig 2.

Ten per cent of subjects randomised to the zero font (no warning) control group mistakenly recalled the warning. We note that the sample proportions at font sizes 8, 10, 14, and 18 point were all significantly greater than the false recall rate (p < 0.0001). Presumably false recall was in operation, in an attenuated form, at font sizes other than zero. It is reasonable to suppose that, if the true recall proportion is \( q \), then the total (true + false) recall proportion is \( p = q + \delta(1 - q) \), where \( \delta \) is the false recall rate for the proportion who cannot truly recall. An estimate for \( \delta \) is provided by the proportion recalling at the zero font, and therefore \( p = q + 0.1(1 - q) \). Thus estimated true recall proportions for font sizes 8, 10, 14, and 18 were 0.51, 0.56, 0.64, and 0.69, respectively—in other words, between 91–96% of total recall. Averaging over contrasts, true and total recall curves are shown in fig 2.

### DISCUSSION

All ads that included a health warning, regardless of font size or contrast, generated significantly greater recall than the ad with no warning. This suggests warning communication has not been negated by exposure to the health warnings required by statute for over a decade. False recall rates were relatively low (10%), and similar to those observed by Popper and Murray when warnings were introduced (5%). Although smokeless tobacco ads were available to the target population, actual exposure may be limited by the few ads in circulation, thus warning communication potential is retained.

Warning size clearly matters, as recall increased significantly with font size. Here 69% recalled the high contrast warning message in 10 point font, and 62% recalled the low contrast warning in the 12 point font—the minimum levels currently mandated for a full page ad. Doubling warning size...
to a 20 point font would increase recall from 63% to 76%, representing a 20% improvement.

Our study provides a model from which recall outcomes can be estimated for each font and warning size, which has implications for attention and retention in health communications that impact attitude and behaviour. The current minimum requirement with 10 point font calls for a symbol that is 1.25 inches (3 cm) wide by 1.42 inches (3.5 cm) high. The rectangular area implied by these dimensions is 1.76 square inches, or 2% of a full page magazine ad.† Treatment groups tested 8 to 18 point font warnings (1–7% of the ad area), from which a continuous range of recall estimates were modelled. As shown in Table 3, a 24 point font warning (12.7% of the ad area) is estimated to yield between 74.9% and 78.3% recall depending on contrast. Generally, we observed a pattern of diminishing returns, especially after 24 point font. Our contrast manipulations yielded negligible effects.

In establishing minimum requirements, a key question is what proportion of those who see a magazine ad for smokeless tobacco should remember the health warning. The model can estimate potential levels of recall for any given warning size, but confidence in recall estimates decreases as sizes grow substantially larger than those observed. Moreover, designers may craft ads to minimise warning prominence and viewer attention will vary. One possible benchmark is warning recall by the proportion who remember the brand name—84% in this study. Table 3 suggests that more than 20% of the ad area would be required. Achieving the ideal goal of warning 100% of the target population would require strategies beyond moderate warning manipulations.

Other findings
High contrast warnings were associated with greater recall, but the difference from low contrast warnings was not significant for proven aided recall in our study or by Popper and Murray. This study demonstrated more substantial and significant effects of warning size on recall, likely due to the larger sample and wider range of font sizes tested than Popper and Murray. Differences in sample composition, test administration, and changes occurring over the intervening decade (exposure) probably contributed as well.

Recall had direct relationships with three demographic variables: education, unemployment (versus employed or in school), and Hispanic ethnicity. The research literature based on undergraduates does not offer relevant material on education or employment status. Those who are employed, in school, or better educated may be more aware of health risks. Personal interview was chosen over self administration to reduce problems associated with literacy, but education and unemployment may be related to recall and recognition task skills.

Compared to non-Hispanic Whites, Hispanics had the lowest recall of the racial/ethnic groups in our sample. Even though recruiters screened for English fluency, the variable “Hispanic” may act as a proxy for other differences that impact message recall, such as English literacy, level of acculturation, or exposure to ads in US magazines.‡

We found no significant relationships of recall with cigarette smoking, smokeless tobacco use, cigarette smoking, (self reported) previous exposure to tobacco product ads, or age,‡ and too few subjects were married to evaluate the influence of marital status which may be a proxy for motivations.

---

†Note that the rectangular area is somewhat greater than the actual amount of space consumed by the irregularly shaped warning symbol.

‡There is a small non-significant relationship between age and recall controlling for education—that is, recall is lower for older subjects when education in the same.

APPENDIX

Two of the advertisements used in the test: high versus low contrast in 10 point font
associated with family responsibilities. This study, as well as Popper and Murray, found no relationship between recall of smokeless tobacco ad warnings and use of either cigarettes or smokeless tobacco. Perhaps related to context (magazine viewing versus tobacco purchasing and using), the recognition phenomenon may operate differently for print ad warnings than for product warnings labels.

This paper focused on proven aided recall (warning recall confirmed by message recognition). Not reported are results for unaided recall, aided recall, and proven unaided recall. The overall frequency of correct responses for proven aided recall was lower than simple aided recall, but higher than unaided recall or proven unaided recall.1,1

Limitations
The demographic composition of this sample of 16–24 year old males encountered at urban and suburban shopping malls in Massachusetts suggests they are reasonably representative of that subpopulation, but generalisation to the larger male subpopulation is inappropriate. Nor can the results be assumed to be similar for females, persons over 24 years old, or residents of rural areas or regions outside Massachusetts. The population of interest is young males targeted by smokeless tobacco advertising, but geographic differences in smokeless tobacco use would be important to consider.

ACKNOWLEDGEMENTS
Support for this research was provided by the Massachusetts Tobacco Control Program.

Authors’ affiliations
C Celebucki, Massachusetts Tobacco Control Program, Massachusetts Department of Public Health, Boston, Massachusetts, USA

REFERENCES