Effects of Pictorial Warning Labels for Cigarettes and Quit-Efficacy on Emotional Responses, Smoking Satisfaction, and Cigarette Consumption

Daniel Romer, PhD 1 · Stuart G. Ferguson, PhD 2 · Andrew A. Strasser, PhD 3 · Abigail T. Evans, PhD 4 · Mary Kate Tompkins, MA 4 · Joseph Macisco, BA 4 · Michael Fardal, MA 4 · Martin Tusler, MS 4 · Ellen Peters, PhD 4

© The Society of Behavioral Medicine 2017

Abstract

Background Experimental research on pictorial warning labels for cigarettes has primarily examined immediate intentions to quit.

Purpose Here, we present the results of a clinical trial testing the impact on smoking during and after a 28-day period of naturalistic exposure to pictorial versus text-only warnings.

Methods Daily cigarette smokers (N = 244) at two sites in the USA were randomly assigned to receive their regular brand of cigarettes for 4 weeks with one of three warnings: (a) text-only, (b) pictures and text as proposed by FDA, or (c) the warnings proposed by FDA with additional text that elaborated on the risks of smoking. Analyses examined the effects of pictorial versus text-only warnings and self-efficacy for quitting on cigarette consumption during and 1 month after the trial as mediated by emotional and cognitive responses as well as satisfaction with smoking.

Results Stronger emotional responses to pictorial than text-only warnings predicted reduced satisfaction with smoking during the trial and lower cigarette consumption at follow-up among the majority of smokers who continued to smoke. Consistent with the efficacy-desire model, those with moderate efficacy reported the greatest reduction in consumption at follow-up. However, a small proportion of smokers (7%) who reported 7-day abstinence at follow-up did not exhibit a significant relation with self-efficacy.

Conclusions Pictorial warning labels proposed by FDA create unfavorable emotional reactions to smoking that predict reduced cigarette use compared to text alone, with even smokers low in self-efficacy exhibiting some reduction. Predictions that low self-efficacy smokers will respond unfavorably to warnings were not supported.

Keywords Warning labels · Emotion · Smoking · Self-efficacy · Smoking satisfaction

Introduction

The Family Smoking Prevention and Tobacco Control Act of 2009 (Tobacco Control Act or TCA) gave the FDA regulatory authority over the marketing of tobacco products and mandated the placement of nine new text warnings on cigarette packs that include a graphic image further illustrating the risks of smoking. However, the introduction of the new program was delayed due to legal action by the tobacco industry, which led the court to conclude that the images selected by FDA mainly evoked emotion, did not contain any relevant information for smokers, and were largely ineffective in discouraging smoking [1].

The conclusions reached by the court were surprising to many researchers because studies had shown that pictorial labels are more noticeable and emotionally engaging than text-only labels [2] and more effective in informing smokers of the risks of smoking [2, 3]. A longitudinal observational
study of the Canadian experience with pictorial warnings, the first country to implement them, found increases in quit attempts and reductions in smoking following implementation [4]. However, observational studies do not rule out other market forces (e.g., changes in promotional activity) that could also affect smokers’ quit intentions and reactions to cigarettes. On the other hand, experimental research that is not susceptible to such confounds has been dominated by one-time exposures to warnings with effects on immediate intentions to quit and related reports [5], leaving open whether pictorial warnings affect smoking behavior.

More recently, researchers have conducted clinical trials in which smokers were randomly exposed to pictorial versus text-only warnings over longer periods. In one large trial conducted in the USA, the pictorial warnings proposed by FDA were compared to the present Surgeon General’s text-only warnings on cigarettes [6]. In that study, a larger proportion of smokers exposed to the pictorial warnings reported quitting smoking during the 28-day trial. However, that study did not isolate the effects of the pictures versus the novel textual information that was mandated by the TCA. Thus, the benefit of adding pictures was not clearly assessed in that study.

In a second study, Evans et al. [7] compared the FDA’s proposed warnings containing both pictures and text against warnings that only contained text. Results indicated that pictorial warnings made smokers feel worse about their habit than warnings containing text alone and that this effect was linked with enhanced risk perceptions and quit intentions. In addition, the negative emotion elicited by pictorial warnings was associated with increased scrutiny of the risk messages and perceived credibility of the warnings, which were also related to quit intentions. However, a third condition that elaborated on the risk beyond that mandated by the TCA, as is done in Canada and other countries, did not enhance the effects of the warnings and actually reduced their credibility, which further reduced effects on intentions to quit.

The Evans et al. study suggested that pictorial warnings may be more effective in communicating the risk of smoking than text alone. However, the analysis did not examine effects on actual smoking behavior. Here, we examine the emotional and cognitive (perceived credibility) effects of warning information on this outcome both during the trial and at a 1-month follow-up. In considering these effects on smoking, we examined an important mechanism by which pictorial warnings could reduce smoking. If such warnings make smokers feel worse about smoking than text alone, they should experience reduced reward from smoking. One way to measure reward value is to assess the satisfaction that smokers receive from their habit. Consistent with this hypothesis, Schüz and Ferguson [8] found that reported satisfaction mediated the effect of a nicotine-patch treatment on reductions in smoking. Following on this finding, we assessed daily reports of satisfaction to test the hypothesis that satisfaction with smoking would mediate the effect of pictorial warnings on smoking behavior.

In addition to examining the effects of pictures and text versus text alone, we were also interested to determine whether smokers with weak self-efficacy to quit would respond as strongly as those with greater self-efficacy. A long line of research suggests that warnings that evoke fearful emotional reactions will be especially ineffective among those who feel unable to change their behavior [8–10], a condition that is likely to characterize life-long smokers. Some researchers have even suggested that strong pictorial warnings will be counterproductive and potentially lead to greater smoking [11, 12].

In a recent model of the effects of self-efficacy for quitting smoking, Romer et al. [13] noted that for an addictive behavior such as smoking, persons with lower self-efficacy actually have greater desire to quit than those higher in self-efficacy. This negative relation results from a history of trying to quit that reduces self-efficacy despite continued and increasing desire to do so. This “efficacy-desire” model suggests that in response to a threatening warning, low self-efficacy smokers will experience stronger unfavorable feelings about their habit than those higher in self-efficacy, an effect that should over time reduce satisfaction with smoking and motivate greater reductions in smoking. However, smokers low in self-efficacy, who are also more addicted to their habit, are likely to experience stronger withdrawal symptoms if they attempt to cut down on smoking, an effect that would maintain the reinforcement value of the behavior [14]. The efficacy-desire model proposes that in predicting behavior change, self-efficacy and desire to quit will combine multiplicatively, with the result that, over time, smokers at intermediate levels of self-efficacy should exhibit the strongest reductions in smoking. In our trial, all three warning conditions provided new and potentially threatening information to smokers. Thus, we tested the efficacy-desire model’s predictions across the three experimental conditions.

To obtain a robust test of the effects of warnings and self-efficacy, we examined smoking both during the trial and at the 30-day follow-up. Cigarettes were provided free of charge during the trial to ensure exposure to the warnings. However, given that tobacco costs are a strong predictor of its use, the full effects of the warnings were not expected until smokers resumed purchasing cigarettes when the negative feelings elicited by the pictorial warnings would be expected to remain in memory [15, 16] but the formerly free cigarettes would no longer serve as a deterrent to quitting.

The proposed causal model in Fig. 1 was used to test the effects of the three warning conditions and self-efficacy on smoking during and after the trial. Hypothesized directions of effects shown in the figure indicate that the previously observed effects of pictorial warnings on emotion and perceived warning credibility would lead to reduced satisfaction
formation that was mandated by the TCA. Thus, the benefit of isolate the effects of the pictures versus the novel textual in-

warnings on cigarettes [6]. In that study, a larger proportion 
effect of a nicotine-patch treatment on reductions in smoking. 

If such warnings make smokers feel 
ed an important mechanism by which pictorial warnings 

follow-up. In considering these effects on smoking, we exam-

mation on this outcome both during the trial and at a 1-month 

which further reduced effects on intentions to quit. 

addition, the negative emotion elicited by pictorial warnings 

related to quit intentions. However, a third condition that elab-

was associated with increased scrutiny of the risk messages 

proposed warnings containing both pictures and text against 

text-only warnings over longer periods. In one large trial con-

ior such as smoking, persons with lower self-efficacy actually 

smoking, Romer et al. [13] noted that for an addictive behav-

[11, 12]. 

The proposed causal model in Fig. 1 was used to test the 
s predictions across the 

intermediate 

costs of smoking returned. Finally, we predicted that low self-

efficacy for quitting should motivate greater smoking reduc-

as mediated by emotional reactions, but ultimately less 

reduction than exhibited by intermediate levels of self-

efficacy.

Method

Participants

The trial was registered on clinicaltrials.gov (NCT01782053) and recruited smokers through fliers and advertisements on social media websites in one mid-western and one east coast US city. As previously reported [7], we screened 1028 persons (consort figure is available in that report) excluding individuals who reported smoking fewer than five cigarettes a day or more than two packs per day, were currently trying to quit smoking, planned on leaving the area within the next 3 months, had a medical condition for which smoking was counter-indicated, who reported inability to read from a computer monitor, or who reported a history of substance abuse (other than smoking or light use of marijuana). Only smokers with low current intentions to quit in the next 30 days were enrolled (scored as rating their intention as a 1 or 2 on a 4-

point scale going from 1 (not at all) to 4 (very likely). At the 

first visit to the research site and all subsequent visits, partic-

ipants were required to provide a carbon monoxide breath sample (Vitalograph, Lenexa, KS) to biochemically verify their smoking status. Participants with less than 5 ppm of expired carbon monoxide (CO) at the first visit were excluded. 

As previously reported [7], we enrolled 293 smokers 

Med = 33.68, SD = 11.56, 45% female), of whom 49 

(17%) did not complete the trial due to inability to adhere to the protocol. In five cases, participants were later discovered to have a medical condition unrelated to the trial that preclud-
ed continued participation. There was no differential attrition during the trial between experimental conditions in regard to baseline smoking rates or important demographic differences, and we confirmed that effects of warnings were robust even with imputation for those missing cases [7]. Thus, the results we present here are based on the 244 participants who com-

pleted the trial. Of these smokers, we were able to contact 174 

(71%) at about 1 month following the trial, and we examined their self-reported daily smoking at that time. Table 1 shows 

that there were no differences in attrition related to experimen-
tional condition or other characteristics in smokers who complet-
ed the 1-month follow-up.

Procedures

Participants were stratified on the basis of gender, amount 

smoked (less than one pack per day vs. more than one pack), 

and quit intentions and then randomly assigned to one of three 

conditions: text only, image (picture plus text), and elaborated 

picture plus text including elaborated information). Over a 

period of 5 weeks, participants came to the lab on a weekly 

basis at which time they completed computer-administered 

surveys about their experience with the new labels and were 
given a 1-week supply of their preferred brand of cigarettes. 
The first week of the trial served as a pretest in which partic-

pants were given cigarettes in non-modified packages based 

on self-reported amount smoked. In the subsequent 4 weeks, 

participants were given cigarettes in packages modified for 

their experimental condition. Participants were required to re-

turn their used filters in plastic pouches, one for each day of 

the week, a procedure developed in previous trials by Strasser et al. [17]. Unused cigarettes were also to be returned. In order 
to avoid the need to buy cigarettes without test warnings, we 

provided participants with an additional pack that if unused 

was also to be returned. Participants were also told that they 
could stop smoking at any time without being dropped from the 

trial.

In addition to returning filters at each weekly lab visit, participants were asked to record every cigarette smoked on 
a smart phone (LG Optimus One, model P500H) modified to 
run study-specific software [18]. Participants were also asked
to complete a report immediately upon waking each day and an evening report between 7 p.m. and midnight each day, to answer randomly timed assessments (average four to five per day), and to complete assessments on a randomly selected subset (target average five per day) of cigarette reports. Those entries were downloaded at each weekly meeting.

All participants received the same nine basic text messages (e.g., “Cigarettes cause fatal lung disease”) mandated by the TCA. Those in the text-only condition received packages modified with only this basic text information, which was placed on the side of each package (for all materials, see Evans et al.). Participants in the pictorial warning conditions received the same basic text but paired with the pictorial warnings proposed by FDA for inclusion on US cigarette packs [19]. Participants in the elaborated text condition received the FDA warnings plus information that described smoking-related risks in greater detail (e.g., “Every cigarette you smoke increases your risk of crippling, often fatal, lung diseases such as emphysema”). Warnings were assigned in a balanced fashion so that participants were exposed to each label in their experimental condition in approximately equal proportions over the trial.

In addition to receiving a supply of study cigarettes at no cost, participants were paid up to $235 based on compliance with experimental procedures over the course of the study. The protocol was approved by the IRB of the University of Pennsylvania, which served as the IRB of record for the trial.

---

Table 1 Demographic and smoking characteristics for those completing the trial versus those reached or not at the 1-month follow-up

<table>
<thead>
<tr>
<th>Variable</th>
<th>Status</th>
<th>Completed trial (N=244)</th>
<th>Completed follow-up (N=174)</th>
<th>Did not complete follow-up (N=70)</th>
<th>P value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (%)</td>
<td>54.5</td>
<td>56.3</td>
<td>43.7</td>
<td>0.523</td>
<td></td>
</tr>
<tr>
<td>Race-ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>58.6</td>
<td>60.9</td>
<td>58.5</td>
<td>0.630</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>32.4</td>
<td>31.0</td>
<td>37.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>3.3</td>
<td>2.9</td>
<td>4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1 reported</td>
<td>3.3</td>
<td>3.4</td>
<td>2.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian and other</td>
<td>2.5</td>
<td>1.7</td>
<td>4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;High School (HS)</td>
<td>8.2</td>
<td>6.3</td>
<td>13.0</td>
<td>0.380</td>
<td></td>
</tr>
<tr>
<td>HS degree</td>
<td>25.1</td>
<td>25.3</td>
<td>24.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>42.4</td>
<td>41.4</td>
<td>44.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associate degree</td>
<td>8.6</td>
<td>9.2</td>
<td>7.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor degree</td>
<td>14.4</td>
<td>16.7</td>
<td>8.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masters or higher</td>
<td>1.2</td>
<td>1.1</td>
<td>1.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menthol smoker (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57.8</td>
<td>57.5</td>
<td>58.6</td>
<td>0.875</td>
<td></td>
</tr>
<tr>
<td>Breath CO @ intake</td>
<td>19.1 (11.4)</td>
<td>19.0 (10.6)</td>
<td>19.5 (13.3)</td>
<td>0.759</td>
<td></td>
</tr>
<tr>
<td>Nicotine dependence (FTND)</td>
<td>4.4 (1.9)</td>
<td>4.3 (1.9)</td>
<td>4.5 (1.8)</td>
<td>0.370</td>
<td></td>
</tr>
<tr>
<td>Reported cigarettes/day @ intake</td>
<td>16.4 (7.5)</td>
<td>16.5 (7.4)</td>
<td>16.1 (7.8)</td>
<td>0.736</td>
<td></td>
</tr>
</tbody>
</table>

Values shown are means (and standard deviations) for continuous variables, and percents for categorical variables. Values by condition are not shown because of the absence of significant differences for all measures.

<sup>a</sup> Determined by F test (for continuous variables) or chi-square (for categorical variables) for difference between those who completed the follow-up versus those who did not.
Measures

Baseline smoking was the average number of filters returned per day for days 2–7 of the pretest week ($M = 15.4$, $SD = 7.14$). This measure was highly related to self-reports of smoking at intake, $r = .83$, $p < .001$. Nicotine dependence was assessed at baseline using the Fagerström Test of Nicotine Dependence [20].

Daily Smoking Rate: We obtained two measures of daily smoking: filters returned and cigarettes logged on the smartphone. For filters returned, we did not include counts for the days on which participants came to the lab (as these did not encompass the entire day). Thus, we analyzed counts for both measures for 6 days per week (6 at baseline and 24 over the exposure period). Compliance for returning filters was high. The median percent of days on which filters were returned was 96.9% across the 5 weeks of the trial. Although participants logged fewer cigarettes per day than filters returned (filters: $M = 15.7$, $SD = 7.55$; logged: $M = 14.8$, $SD = 7.10$; $t = 5.21$; $p < .001$), the two measures were highly correlated, $r = .93$. We report results based on filters returned, with a separate test of robustness using logged cigarettes.

Smoking at the follow-up was obtained in a phone interview in which participants were asked first if they had smoked in the past 7 days, and if so, how many cigarettes they had smoked per day during that period. We obtained those reports from all but three of those interviewed.

Smoking satisfaction was collected during the daily “evening report” when participants reported how satisfying smoking was during that day, measured on a 0–100 continuous scale from “No!!” to “Yes!!” [8]. The vast majority of participants (96.3%) reported their evening satisfaction on at least half of the days during the 28-day exposure period ($M = 24.2$ days), with a mean decrease in reporting of only .6 days from the first to the second half of the exposure period.

Emotion regarding smoking was assessed three times over the trial (after 1, 2, and 3 weeks of exposure to the labels) in response to the question: “Did the packaging change how you feel about smoking?” with responses ranging from “the packaging made me feel much worse about smoking” (1) to “… much better about smoking” (7) [21]. We reverse scored the scales and used each assessment as an indicator of a latent variable of negative emotion toward smoking. These negative reactions increased only slightly over the course of the trial (from 5.0 to 5.2) and the time change did not interact with label condition ($p = .52$), suggesting that scores at each wave were adequate for capturing this effect. Another form of emotion, agitation, was measured by having participants view and rate each warning label that had been on their cigarette packages after the first 2 weeks as to whether or not “the information makes you feel agitated,” on a scale from (1) not at all agitated to (7) very agitated. An agitation score averaged across labels (Cronbach’s alpha ($\alpha$) = .93) was moderately correlated with ratings of feeling worse about smoking ($r$’s from .42 to .52) and, thus, was included as an indicator of the latent emotion factor, with excellent fit: $X^2(2) = 1.28$, $p = .53$, RMSEA = .00, and CFI = 1.00.

A measure of warning credibility was obtained after 2 and 4 weeks of exposure to the warnings at which time participants rated the extent to which they believed each warning on a scale from “completely not believable” (1) to “completely believable” (7) [22]. These ratings had high reliability ($\alpha$’s = .89 and .92), with means that did not change over the course of the trial ($M = 6.2$ at both times). Thus, they were used as indicators of a latent credibility factor that served as the cognitive mediator of warning label effects.

Self-efficacy for quitting was assessed after the baseline period of the trial, with two items that have been shown to predict responses to warning labels [13]: “It is hard for a smoker to quit smoking” (reversed) and “I would not need help from anyone to quit smoking” answered on a 1 to 7 disagree-agree scale. These items represent agreement with both general knowledge that quitting is difficult and the recognition that it would also be difficult for the responding smoker to quit. The items were correlated ($r = .32$, $p < .001$), and as in previous research [13], the mean of these items served as the measure of self-efficacy. To validate that self-efficacy was inversely related to desire to quit, we assessed this desire on the same occasion as self-efficacy with the mean of three items: “Do you want to quit smoking?”), “Do you wish you did not smoke?”, and “If you could easily quit right now, would you?”, all answered on a 4-point scale from not at all (1) to a lot (4). This score ($\alpha = .86$) was inversely related to self-efficacy, $r = -.19$, $p = .002$. Self-efficacy was unrelated to experimental condition at the outset of the trial, $F(2241) = .82$, $p = .44$, and remained so during the remaining weeks of the trial, all $p$’s > .36.

Preliminary examination of the relation between self-efficacy and smoking at follow-up indicated a quadratic trend that could be driven by smokers at the highest level of efficacy (see Fig. 2a). Because of the small number of cases at that end of the scale ($n = 2$), we truncated self-efficacy scores by recoding these two scores into the next lowest score. This retained those cases while reducing their extremity effects. To assess the hypothesis that smokers at intermediate levels of self-efficacy would reduce their smoking the most, we created a quadratic contrast by squaring deviations from the middle of the scale. Thus, higher numbers indicate smokers with lower or higher levels of self-efficacy. To contrast this quadratic trend with the relation between self-efficacy and emotion, we show this relation in Fig. 2b, which clearly shows a linear relation.
Data Analysis

We conducted preliminary analyses of trends in daily satisfaction and smoking over 28 days of the trial using mixed-effects models implemented in Stata [23]. Based on those analyses, we reduced daily rates of satisfaction and smoking to the mean of each score over the course of the trial as described in the “Results.” As in the first report from this trial, we studied two planned contrasts: one contrasting the two picture conditions (+.5) with the text-only condition (−1) and a second orthogonal contrast comparing the elaborated condition (−1) versus the image condition (1), with the text-only condition coded 0. Structural equation analysis (SEM) of the model in Fig. 1 was conducted with MPlus [24], with model fit evaluated using multiple global fit indices. Significance levels and 95% confidence limits (CLs) were obtained using bias-corrected bootstrapping procedures [25].

As described in the “Results,” some smokers reported not smoking in the past 7 days at the follow-up. Because this produced a surplus of zeros in the response distribution that was not present during the trial, we employed a zero-inflated model to study both the continuous smoking reports and reports of not smoking. However, because this analysis does not produce overall goodness of fit statistics nor permit tests of indirect effects, we also tested each outcome separately using bootstrapping procedures [25].

Although there were very few missing data during the trial, follow-up data were missing in about 30% of the 244 cases. Because those missing scores were not differentially related to experimental conditions or in regard to smoking rates, nicotine dependence, or major demographic variables (see Table 1), we employed the MPlus procedure of full information maximum likelihood estimation to impute scores for missing cases. This procedure uses the correlation matrix for the full sample to predict outcomes of the missing cases at follow-up, a procedure recommended with data that are missing at random [26].

Results

We began by examining trends in daily satisfaction and smoking controlling for baseline smoking. There were no differences between conditions in baseline satisfaction, $F(2,237)=1.48, p = .23$. However, preliminary analysis using mixed-effects regression indicated that satisfaction declined linearly over the course of the trial in all three conditions ($b = -.21, z = -4.35, p < .001$), with no interaction with experimental conditions ($b's = -.14, -.02; z's = -1.04, -.11; p's > .25$). There were also no differences in intercepts during the exposure period ($b's = -2.3, -3.7; z's = -8.0, -1.31; p's > .15$). Baseline satisfaction did not correlate with mean satisfaction over the exposure period, $r = .01$. Both baseline smoking ($b = 4.1, z = 2.20, p = .027$) and nicotine dependence ($b = 1.50, z = 2.18, p = .029$) were related to satisfaction. However, neither demographic differences (age, gender) nor self-efficacy predicted satisfaction.

There were no baseline differences between conditions for either filters returned, $F(2, 241) = 0.79, p = .45$, or cigarettes logged, $F(2, 239) = 1.03, p = .36$. Raw trends in filters returned over the period of exposure to the new labels are in the Electronic Supplement. On average, smoking was stable over the course of the trial, $b = .017, z = 1.46, p = .15$, and did not differ by condition, $b's = -.51, -.16; z's = -1.25, -.39; p's > .20$. There were also no differences by age ($b = .014, z = .89, p = .33$), gender ($b = .29, z = .82, p = .42$), linear self-
efficacy ($b = -.02, z = -.07, p = .94$), or quadratic self-efficacy ($b = -.24, z = -1.12, p = .26$). Nevertheless, filters returned were related to baseline smoking ($b = .96, z = 37.1, p < .001$) and nicotine dependence ($b = .27, z = 2.76, p = .006$).

Given that the slopes in satisfaction and smoking rates did not differ by condition over the course of the trial, we used each participant’s mean satisfaction and smoking rate across the 4-week period in further analyses. The correlation matrix for all variables used in the analysis is in the supplement. Baseline smoking and nicotine dependence were included in the SEM to control for the significant relations between these individual differences and both satisfaction and smoking during and after the trial. Participants with higher levels of smoking complied more often with reports of satisfaction ($r = .26, p < .001$), did so more often in the image conditions ($r = .15, p = .01$), and reported higher satisfaction ($r = .11, p = .08$). A preliminary analysis controlling for compliance revealed that the findings were unaffected by this reporting difference.

**Follow-up Reports of Smoking**

The average time from the end of the trial to the follow-up was 34.8 days (SD = 10.5). Time to the follow-up interview was uncorrelated with reports of smoking at that time ($r = -.028$), and there was no relation between the date of the follow-up and smoking, controlling for baseline smoking ($p = .30$). Twelve participants (7% of those in the follow-up) reported not smoking in the prior 7 days. They tended to be lighter smokers during the trial than other participants ($M = 10.1$ vs. $15.9, p = .01$). Although smoking during the trial did not differ from baseline as assessed by filters returned (15.4 vs. 15.7), smoking reports among all assessed at follow-up ($M = 12.4$) were lower than self-reported smoking as measured at intake ($M = 16.4$) and filters returned during the trial ($M = 15.7$). However, there were no differences between conditions in reported smoking at follow-up, either for those who reported smoking in the past 7 days, $F(2,156) = .37, p = .69$, or for those who reported not smoking during that time, $X^2(2) = 2.31, p = .32$.

**Tests of the Model in Fig. 1**

Initial tests of the SEM using a zero-inflated model revealed that neither contrast for self-efficacy predicted smoking during the trial ($p’s = .65, .16$) and the quadratic contrast did not directly predict emotion ($p = .79$). Nicotine dependence also did not predict emotion ($p = .34$). In addition, credibility, which was lower in the elaborated condition, did not predict satisfaction ($p = .91$) nor smoking during the trial ($p = .50$) and was uncorrelated with smoking at the follow-up ($r = .07$). Finally, emotion did not predict abstinence at the follow-up ($p = .25$). After these preliminary analyses, Fig. 3 shows the resulting standardized weights, with unstandardized results in Table 2.

Because the zero-inflated model did not permit a test of model fit or of mediation, we also tested separate models for the continuous and binary abstinence outcomes. The results of those tests, also shown in Table 2, confirm excellent fits for each outcome. The model explained 14.1% of the variance in emotion, 9.3% of the variance in satisfaction, 89% of the variance in filters returned, 87% of the variance in follow-up reports of smoking, and 65% of the variance in abstinence. In describing the results, we use the zero-inflated model for direct effects shown in Table 2 and the separate models for indirect effects on smoking.

As previously reported [7], the image conditions produced more negative emotion regarding smoking, $b = .32, p < .001$. Here, we find that this affective reaction was associated with reduced satisfaction with smoking, $b = -.45, p = .002$, which in turn was linked to change in daily smoking rates, $b = .022, p = .016$. The mediated path from the image conditions (vs. text alone) through negative emotion and smoking satisfaction to cigarettes consumed during the trial predicted a small reduction of .024 (95% CI = −.003, −.073) cigarettes per day.

Smokers with higher self-efficacy experienced a weaker emotional response than low self-efficacy smokers, $b = -1.16, p < .001$, and the mediated path from linear self-efficacy through emotion and satisfaction to smoking during the trial predicted a small relation, $b = .012 (95% CI = .002, .039)$. In addition, low self-efficacy smokers experienced directionally more satisfaction than higher self-efficacy smokers, $b = -.90, p = .50$; although not significant, this finding was in the opposite direction of what would be expected based on the low
Table 2  Results of SEM with unstandardized parameters and robust confidence intervals (CIs) for zero-inflated model and bias-corrected bootstrapped CIs for separate models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Zero-inflated model (N = 244)</th>
<th>Continued smoking model (N = 232)</th>
<th>Smoking abstinence model (N = 244)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate 95% confidence interval</td>
<td>Estimate 95% confidence interval</td>
<td>Estimate 95% confidence interval</td>
</tr>
<tr>
<td>Emotion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling week 1</td>
<td>1.000 1.026, 1.374</td>
<td>1.000 1.043, 1.429</td>
<td>1.000 1.060, 1.448</td>
</tr>
<tr>
<td>Feeling week 2</td>
<td>1.214 1.014, 1.414</td>
<td>1.214 1.041, 1.487</td>
<td>1.221 1.044, 1.495</td>
</tr>
<tr>
<td>Feeling week 3</td>
<td>1.096 0.814, 1.378</td>
<td>1.129 0.856, 1.486</td>
<td>1.110 0.858, 1.444</td>
</tr>
<tr>
<td>Agitated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 2</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 4</td>
<td>1.119 0.819, 1.419</td>
<td>1.068 0.785, 1.462</td>
<td>1.110 0.862, 1.525</td>
</tr>
<tr>
<td>Picture versus text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion</td>
<td>0.320 0.174, 0.488</td>
<td>0.312 0.153, 0.486</td>
<td>0.315 0.155, 0.491</td>
</tr>
<tr>
<td>Emotion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td>-0.230 -0.364, -0.096</td>
<td>-0.240 -0.372, -0.108</td>
<td>-0.229 -0.361, -0.097</td>
</tr>
<tr>
<td>Satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial smoking</td>
<td>0.022 0.004, 0.040</td>
<td>0.017 0.001, 0.036</td>
<td>0.022 0.005, 0.040</td>
</tr>
<tr>
<td>Follow-up smoking</td>
<td>0.781 0.643, 0.919</td>
<td>0.749 0.606, 0.881</td>
<td>-</td>
</tr>
<tr>
<td>Follow-up abstinence</td>
<td>-0.101 -0.277, 0.075</td>
<td>-</td>
<td>-0.008 -0.014, -0.002</td>
</tr>
<tr>
<td>Linear self-efficacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion</td>
<td>-0.155 -0.237, -0.073</td>
<td>-0.163 -0.249, -0.072</td>
<td>-0.158 -0.242, -0.074</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>-0.904 -3.558, 1.750</td>
<td>-1.637 -4.514, 1.944</td>
<td>-1.104 -3.308, 1.600</td>
</tr>
<tr>
<td>Follow-up smoking</td>
<td>0.505 -0.215, 1.225</td>
<td>0.372 -0.686, 1.182</td>
<td>-</td>
</tr>
<tr>
<td>Follow-up abstinence</td>
<td>0.544 -0.214, 1.302</td>
<td>-</td>
<td>0.044 -0.021, 0.121</td>
</tr>
<tr>
<td>Quadratic self-efficacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up smoking</td>
<td>0.447 0.105, 0.789</td>
<td>0.413 0.040, 0.787</td>
<td>-</td>
</tr>
<tr>
<td>Follow-up abstinence</td>
<td>0.723 -0.035, 1.481</td>
<td>-</td>
<td>0.027 0.001, 0.055</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$ (76)</td>
<td>114.28, p = .003</td>
<td>112.16, p = .004</td>
<td>112.16, p = .004</td>
</tr>
<tr>
<td>CFI = .97 TLI = .97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEA = .047 (90% CI = .028, .063)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

self-efficacy smokers’ greater negative emotional response. In total, there was no relation between linear self-efficacy and smoking during the trial, $b = -0.015$ (95% CI = -0.104, 0.031).

After controlling for smoking during trial, the more emotionally upsetting the warning during the trial, the less smoking was reported at the follow-up, $b = -1.53$, $p = 0.002$. As a result, the pathway from the image conditions to follow-up smoking mediated by emotion predicted a reduction of .47 cigarettes per day (95% CI = -0.97, -0.17). Satisfaction with smoking during the trial did not directly predict follow-up reports of smoking, and the effect of the image conditions on smoking at the follow-up as mediated by satisfaction was not significant, $b = -0.018$ (95% CI = -0.056, -0.003). In total, however, the two paths from the image conditions to smoking at follow-up predicted a reduction of .48 cigarettes per day (95% CI = -0.99, -0.18), or a 3.3% reduction from the average of 15 per day during the trial.

Linear self-efficacy was positively related to smoking at the follow-up as mediated by emotion, with a total indirect effect of $b = .24$ (95% CI = 0.08, 0.33). Nevertheless, there was no direct effect of linear self-efficacy on smoking at the follow-up, $b = .37$, $p = .377$, and the total linear relation was non-significant, $b = .56$ (95% CI = -0.49, 1.42). However, quadratic self-efficacy directly predicted smoking at the follow-up,
In the prediction of abstinence (vs. smoking) at follow-up, there was a non-significant relation between quadratic self-efficacy and abstinence, \( b = .72, p = .057 \), indicating a trend toward smokers with intermediate self-efficacy reporting less abstinence. There was no indirect effect of the image conditions on abstinence at follow-up, \( b = -.008 \) (95% CI = -.030, .007), and no effect of linear self-efficacy on abstinence, \( b = .05 \) (95% CI = -.02, .13).

To determine the robustness of the model, we tested the model using logged cigarettes during the trial. The results indicated the same pattern of findings regarding the effects of images versus text and both linear and quadratic self-efficacy.

**Discussion**

This research presents one of the first randomized trials to explore the impact of repeated exposure to pictorial warning labels on smoking behavior in a naturalistic setting. Although we did not observe direct effects of pictorial warnings on smoking behavior, the findings indicate that warning labels with text messages mandated by the TCA and the images proposed by FDA increase emotional reactions that predict reduced satisfaction with smoking. Satisfaction with smoking declined in all three experimental conditions; however, reductions in smoking attributable to the emotional effects of the warnings were stronger in the image conditions than in the text-only condition. Furthermore, there was no evidence that the effects of the image conditions on emotion or satisfaction plateaued over the trial, despite the possibility that emotional reactions might subside with time.

In addition to effects of exposure to pictorial labels during the trial, we also observed effects at the 1-month follow-up when participants resumed purchasing their cigarettes. A small proportion of smokers (7%) reported not smoking in the past week, a reduction mainly observed among lighter smokers across all three conditions. Among those still smoking, we observed a nearly half cigarette per day reduction (.48) in reported smoking attributable to pictorial warnings as mediated by the effects of those warnings on emotion. Emotions, as has been found previously, act not only as information but also as direct motivators of behaviors and behavioral intentions [7, 27]. Although reduction in smoking during the trial was partially mediated by effects on satisfaction, the reduction at follow-up was only predicted by the change in emotion brought about by the pictorial warnings during the trial. Thus, these warnings have the potential not only to increase intentions to quit as observed in our first report from this trial [7] but also to reduce smoking rates, an effect that could facilitate subsequent success at quitting [28]. These findings suggest that the conclusion reached by the court that the emotions evoked by pictorial warnings provide no public health benefit is unjustified.

It is likely that the full effects of the warnings were not evident until smokers resumed purchasing cigarettes because this deterrent to smoking (its cost) was much lower during the trial than after, whereas emotional effects of the warnings likely persisted. We did not measure satisfaction with smoking at the follow-up, but it is likely that satisfaction at that time was even lower in the image conditions, an effect that was registered in the direct effect of emotion on smoking at the follow-up. Nevertheless, future research should examine satisfaction with smoking in response to warning labels to determine their effects on this important mediator.

Surprisingly, the effects of the image conditions on the perceived credibility of the warnings did not affect satisfaction with smoking during the trial and did not predict smoking behavior at the follow-up. Furthermore, the adverse effects of the elaborated text condition on credibility reported by Evans et al. [7] did not carry over to either satisfaction with smoking or smoking at the follow-up. This suggests that (1) those cognitive responses were not as powerful as the emotional effects produced by the pictorial warnings and (2) despite the apparent inhibiting effect of elaborated text on reports of warning credibility, the lasting emotional effects of the images dominated smoking behavior. Nevertheless, it is still possible that other cognitive responses, such as self-generated beliefs about the benefits of quitting, that were not measured in this study could have affected smoking. Future research should also explore the opportunity to develop elaborated text that can further enhance the emotional effects of pictorial warnings [22]. Such text has been used in Canada and other countries, but more research is needed to determine if it adds effectiveness beyond base text and powerful images.

The results regarding self-efficacy for quitting were largely supportive of the efficacy-desire model. Among smokers who continued to smoke, those intermediate in self-efficacy exhibited the greatest reduction in smoking at the follow-up. This outcome is explained by the opposing forces of greater desire to quit among low self-efficacy smokers versus the greater perceived ability to quit among high self-efficacy smokers. High self-efficacy smokers feel less threatened by warnings and less motivated to reduce their smoking. Because these opposing forces combine multiplicatively, smokers at intermediate levels of self-efficacy exhibit the greatest change.

It is important to note that irrespective of self-efficacy, smokers reported smoking less at the follow-up compared to the beginning of the trial, indicating that stronger warnings may be beneficial even for smokers low in self-efficacy. This reduction could not be explained by secular trends in smoking. Among those who reported 7-day abstinence, there was a trend for smokers with intermediate levels of self-
efficacy to be less likely to report this outcome. However, this trend was based on a small sample, suggesting the need for further research with larger samples. Nevertheless, we found no support for the prediction that low self-efficacy smokers will increase their smoking in response to warning labels, as some have suggested [12].

The finding that warnings motivate low self-efficacy smokers to reduce their consumption is contrary to theories of fear appeals [10]. Nevertheless, it should be noted that the prediction that persons with low self-efficacy will exhibit less change than those with greater self-efficacy has not been consistently supported [29, 30], and effects regarding smoking cessation have been weak [31]. Depending on the range of self-efficacy observed, one can find that persons low in self-efficacy report more of the health-compromising behavior than others. However, one can also find no difference if the nonmonotonic relation is not assessed. The efficacy-desire model suggests that for addictive behaviors, more care will be needed to examine relations with self-efficacy.

Although low self-efficacy smokers were more emotionally upset by novel warnings (including those in the text-only condition), they also exhibited a trend toward greater satisfaction with smoking during the trial. This pattern might reflect the difficulty these smokers have in trying to cut down on their habit. One of the characteristics of drug addiction is the tendency for the habit to be maintained by negative reinforcement. When experiencing withdrawal, persons with a drug addiction resume the habit in order to reduce the stress of withdrawal symptoms [14]. As a result, the habit is maintained despite the desire to quit. This could be another manifestation of the difficult situation faced by the smoker with low self-efficacy. Future research should examine different message strategies that might help low efficacy smoker to cope with symptoms of withdrawal.

Although the mediated effect of pictorial labels on smoking at the follow-up was small (about a 3.3% reduction in cigarettes per day), this must be considered in the context of over 40 million smokers in the USA alone who consume an average of 14 cigarettes a day [32]. This reduction in daily smoking would reduce daily consumption by approximately 18.5 million units compared to text-only warnings. In addition, the potential for light smokers to abstain completely at least for some periods would add to these reductions. With continued reductions, persons addicted to smoking might find it easier eventually to abstain from the habit [28].

Our findings should also be placed in the context of stronger warnings than are currently in the US market. The text-only control condition had novel warnings with more information than those currently in use. Thus, it would be expected that smokers would experience reduced satisfaction even in response to these text warnings. Introducing both new text warning information and pictorial content as mandated by the TCA should have an even greater impact compared to current warnings than was observed here. Additionally, our pictorial condition used the nine images proposed by FDA for placement on US cigarette packages in their 2011 Final Rule. Although these warnings elicit more emotional reactions than text-only warnings, they are less evocative than the warnings used in other countries [33]. Past research suggests that placing more emotionally evocative images on cigarette packaging can increase warning impact, which may lead to greater changes in smoking behavior than we observed [8, 33]. A recent analysis suggests that a 5% reduction in prevalence can be realized with the introduction of stronger pictorial labels, such as those used in other countries [34].

One limitation of the present design was the provision of free cigarettes to participants. This choice probably worked against observing larger reductions in smoking during the trial. Consistent with this expectation, we found greater reduction in smoking rates in the image conditions at the follow-up when smokers resumed buying their own cigarettes. This result occurred despite no longer being exposed to the warnings. Unfortunately, our experiment did not permit an estimate of effects on smoking with continued exposure to pictorial warnings. Nor did it provide a test under the more realistic conditions when cigarettes are universally sold with pictorial warnings.

Another limitation was reliance on self-report at the follow-up. Although we could not directly observe reduction in smoking attributable to the image conditions, we have no reason to believe that participants’ reports were biased in favor of those conditions. In addition, in other research, stronger images have been found to produce stronger brain responses in regions that register emotional effects of stimuli and greater response in a region that mediates memory [16]. This is consistent with the findings we observed in the follow-up in that the emotional effects of the images persisted despite the sometimes countervailing effects of cognitive reactions concerning message credibility in the elaborated text condition.

Our requirement that smokers record their daily reactions, including satisfaction with smoking, may raise concerns about reactivity to these assessments. However, considerable experience with such measures suggests that they do not produce results that interfere with what would otherwise appear without them [35]. Furthermore, smokers in all three conditions received modified warnings and so reactivity specific to pictorial warnings seems unlikely.

Although smokers were exposed to new warnings in all conditions, it is also the case that their exposure depended on their rate of smoking. Nevertheless, this rate, on average, was relatively constant across conditions and thus could not explain differential emotional reactions by condition.
Finally, we did not observe direct effects of pictorial warnings versus text-only on smoking either during or after the trial. It was only when we examined the indirect effects produced by those warnings that we were able to observe effects on smoking. The pattern of indirect effects we observed could be used to estimate the power required to observe direct effects in a larger trial with greater statistical power. At the same time, our findings show that smoking was reduced in all conditions after the 28-day exposure period and that even smokers low in self-efficacy were able to reduce their consumption of this harmful product.

Acknowledgements We thank Michael Hennessy, Peter Yang, Lydia Emery, and Kaitlin Sheerin for their assistance with the conduct and analysis of the study. The research was supported by grants from the National Cancer Institute and the Family Smoking Prevention and Tobacco Control Act (R01CA157824) and the National Institutes of Health and the Food and Drug Administration (P50CA180908). However, the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the Food and Drug Administration.

Compliance with Ethical Standards

Authors’ Statement of Conflict of Interest and Adherence to Ethical Standards Authors Romer, Ferguson, Strasser, Evans, Tompkins, Macisco, Fardal, Tusler, and Peters declare that they have no conflict of interest. All procedures, including the informed consent process, were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

References


