Health Psychology

Manuscript version of

Numeracy and Memory for Risk Probabilities and Risk Outcomes Depicted on Cigarette Warning Labels

Brittany Shoots-Reinhard, Breann Erford, Daniel Romer, Abigail T. Evans, Abigail Shoben, Elizabeth G. Klein, Ellen Peters

Funded by:

- Food and Drug Administration, Center for Tobacco Products
- National Institutes of Health, National Cancer Institute
- National Science Foundation

Numeracy and memory for risk probabilities and risk outcomes depicted on cigarette warning labels

Brittany Shoots-Reinhard, PhD[1]*, Breann Erford, PhD[1], Dan Romer, PhD [2], Abigail T. Evans, PhD[1,3,5], Abigail Shoben, PhD[3], Elizabeth Klein, PhD[3], & Ellen Peters PhD[1,4]

1 Department of Psychology, Ohio State University, Columbus, Ohio, United States
2 Annenberg Public Policy Center, University of Pennsylvania, Philadelphia, PA, United States
3 College of Public Health, Ohio State University, Columbus, OH, United States
4 School of Journalism and Communication, University of Oregon, Eugene, OR, United States
5 Battelle Memorial Institute, Health Research, Baltimore, MD, United States

* Corresponding Author: shoots-reinhard.1@osu.edu, 1835 Neil Ave, Columbus, OH 43210

Acknowledgements: Research reported in this publication was supported by grant number P50CA180908 from the National Cancer Institute of the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) Center for Tobacco Products and grant number 1558230 from the National Science Foundation (NSF). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, FDA, or NSF. We thank the OSU-CERTS External Advisory Board and three anonymous reviewers for comments and suggestions.

Numeracy and memory for risk probabilities and risk outcomes depicted on warning labels

Objective. Greater numeracy is associated with higher likelihood to quit smoking. We examined whether numeracy supports learning of numeric health-risk information and, in turn, greater risk perceptions and quit intentions. ***Method***. Adult smokers (N=696) viewed text warnings with numeric risk information four times each in one of three warning-label types (Text-only, Low-emotion pictorial (i.e., with image), High-emotion pictorial). They completed posttest measures immediately or six weeks later. Emotional reactions to warnings were reported the second time participants viewed the warnings. Numeracy, memory for risk probabilities and risk outcomes, risk perceptions, and quit intentions were assessed post-exposures. ***Results***. Memory for risk probabilities and risk outcomes depended on warning-label type and posttest timing. Consistent with memory-consolidation theory, memory for High- vs. Low-emotion labels was lower immediately, but declined less for High- than Low-emotion labels. Label memory was similar between conditions at six weeks. Numeracy predicted overall superior memory (especially for risk probabilities) controlling for health literacy and education. It also indirectly predicted greater risk perceptions and quit intentions via memory. In exploratory analyses, however, the more numerate's superior recall of smokers' risk probabilities was associated with *lower* risk perceptions. ***Conclusions***. Numeracy is associated with superior risk memory, which relates to greater risk perceptions and quit intentions. More numerate and educated smokers may be better able to quit due to their superior learning of smoking's risks.

Keywords: numeracy, memory, emotions, risk perceptions, tobacco control

ClinicalTrials.gov Identifier: NCT03501472

Introduction

Smoking in the United States has declined in recent decades. Although it remains the leading cause of preventable death in the US with more than 480,000 smoking-related deaths per year, only 14% of the US population smoked in 2017 (Wang et al., 2018). However, significant education-related smoking disparities exist such that more (vs. less) educated individuals are less likely to smoke and, if they do smoke, they are more likely to quit (e.g., USDHHS, 2014; Wang et al., 2018; Zhuang, Gamst, Cummins, Wolfson, & Zhu, 2015). Demographic variables (e.g., age, income, insurance status) could explain education-related disparities in smoking and other health behaviors, but studies control for these variables (e.g., Cutler & Lleras-Muney, 2010; Siahpush, McNeill, Borland & Fong, 2006). Education-related smoking disparities may be due to more educated individuals having superior knowledge (e.g., about which diseases are caused by smoking and effective ways to reduce cancer risk; Lillard, 2017). However, cognitive ability explains a larger portion of education's effect on healthy behaviors than health knowledge and demographics; cognitive ability also helps explain health knowledge itself (Peters, Baker, Dieckmann, Leon, & Collins, 2010). Furthermore, education appears to increase cognitive abilities (e.g., Cutler & Lleras-Muney, 2010; Peters et al., 2010). In the present paper, we examine possible effects of one cognitive ability – numeracy (i.e., the ability to understand and use probabilistic and other mathematical concepts) – that may underlie education-related smoking disparities.

Numeracy supports understanding and acquisition of health knowledge and increases its impact on judgments and behaviors (Reyna, Nelson, Han, & Dieckmann, 2009). In fact, when presented with numeric information, such as the probability of risk outcomes, the highly numerate are more likely than the less numerate to pay attention (Fleig, Meiser, Ettlin, &

Rummel, 2017; Jasper, Bhattacharya, & Corser, 2017), derive affective meaning from numeric information (e.g., Peters, Lipkus, & Diefenbach, 2006; Petrova, Plight, & Garcia-Retamero, 2014), remember numbers (Galesic & Garcia-Retamero, 2011), and use them in judgments and decisions (Jasper et al., 2017; Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010). Ultimately, becoming more numerate produces more healthy behaviors (Peters et al., 2017), and the highly numerate enjoy healthier outcomes than the less numerate (e.g., Reyna et al., 2009).

Only two known studies have linked numeracy with smoking. First, greater numeracy is associated with greater smoking cessation: In one study, every one-point increase in numeracy skills corresponded to a 24% increase in the odds of quitting (controlling for reading literacy, education, race/ethnicity, gender, income, and age; Martin et al., 2012). Second, the use of pictorial warning labels (WLs) on cigarette packages may assist the less numerate, in particular, to understand smoking's risks and increase quit intentions (Klein, Quisenberry, Shoben, Romer, & Peters, 2019).

In contrast, considerable research has examined effects of pictorial WLs on smoking-related outcomes (e.g., Noar et al., 2016). Pictorial warnings on cigarette packages elicit greater attention to, emotional reactions to, and scrutiny of the warnings; they also increase long-term risk knowledge (Evans et al., 2015; Hammond, 2011; Noar et al., 2017; Peters et al., 2019). This prior research, however, has focused on pictorial WLs depicting smoking-related negative health outcomes (e.g., smoking can cause cancer), ignoring the likelihoods of those outcomes.

Smokers know little about these probabilities and frequently underestimate them (e.g., Ayanian & Cleary, 1999; Dillard, McCaul, & Klein, 2006; Weinstein, 1998) or overestimate them (e.g., lung cancer risk, Kristiansen, Harding, & Eiser, 1983; Viscusi, 1990). However, informing people about probabilities improves risk comprehension and increases willingness to follow a doctor's advice among more and less numerate individuals (Peters, Hart, Tusler, & Fraenkel, 2014; Schwartz et al., 2017). In the current research, we focus on learning risk probabilities (e.g., the likelihood of cancer from smoking).

In the present study, participants viewed typical warning text that included risk outcomes (e.g., "Cigarettes cause fatal lung disease") in one of three warning-label types (Text-only, Low-emotion pictorial (i.e., with image), High-emotion pictorial). Each warning text also was augmented with risk-probability information for that outcome (e.g., "19.8% of smokers will die from lung disease, compared to 2.0% of non-smokers"). Memory was measured immediately or six weeks after the last warning-label exposure. We measured memory in four ways – recall of smokers' risk probabilities (e.g., "75.4% of smokers"), recall of nonsmokers' risk probabilities (e.g., "21.0% of nonsmokers"), recognition of risk probabilities, and recall of risk outcomes (e.g., "die before the age of 85"). The probabilities, of course, must refer to the risk outcomes. Thus, outcomes served as cues for recall and recognition of risk probabilities (e.g., the cue for the probability 75.4% is "What percent of smokers will die before the age of 85?").

We had three major hypotheses. First, memory-consolidation research has revealed that long-term memories do not form immediately and especially when emotional arousal is involved (Kleinsmith & Kaplan, 1963; Mather, 2007). Instead, high negative and positive arousal produce greater amygdala activation which interacts with other brain regions, such as the hippocampus, to produce greater long-term memory consolidation (Cahill & McGaugh, 1998) up to at least six weeks later (Steidl, Fazik, & Anderson, 2011). In recent research, we demonstrated this memory effect for smoking-related health outcomes, but no information was presented concerning risk probabilities (Peters et al., 2019). As in prior research, we focus on the negative arousal (hereafter referred to as emotional reaction) elicited by the health warnings. We hypothesized:

Hypothesis 1 (H1): Smoking-related health risk information (i.e., risk probabilities and their associated risk outcomes) paired with High-emotion pictorial WLs (vs. lower emotion, including Text-only, warnings) would be recalled less well immediately, but would be retained better over time.

Of course, memory for numeric information is likely influenced by more than emotional reactions. The more numerate recall numeric information more accurately than the less numerate (Galesic & Garcia-Retamero, 2011). Health literacy also has been associated with better memory for smoking risk information (Hoover et al., 2018). Similar to prior research (Martin et al., 2012), we expected independent effects of literacy and numeracy on memory for smoking-related health risk information (i.e., risk probabilities and risk outcomes).

Hypothesis 2 (H2): Those higher (vs. lower) in numeracy would have better memory for risk probabilities and outcomes (controlling for health literacy and demographics).

Risk perceptions and quit intentions also were used to assess whether greater knowledge of smoking risk probabilities and risk outcomes was meaningful to variables crucial to policy makers' ability to design effective future WLs. In particular, we tested whether numeracy correlations with risk perceptions and quit intentions existed and were mediated by greater risk memory (for both probabilities and outcomes). Because emotional reactions also influence memory (H1) and lower numeracy is associated with more negative emotional reactions to math (Peters & Bjalkebring, 2015), we controlled for emotion's known effects on memory, risk perceptions, and quit intentions (Evans et al., 2015, 2017; Peters et al., 2019).

Hypothesis 3 (H3): Better memory for risk probabilities and outcomes would predict greater risk perceptions and, in turn, quit intentions. This superior memory would mediate numeracy's effects on risk perceptions and quit intentions.

Although we expected that more accurate memory would generally predict greater risk perceptions and quit intentions, we also explored the relative predictive power of the four memory indices for risk perceptions and quit intentions. For example, it seemed plausible that memory for smoker-related risk outcomes and probabilities would predict risk perceptions better than that for nonsmoker-related risk probabilities because smoker risk information would be more self-relevant for this population.

Exploratory Hypothesis 3A (H3A): Numeracy would increase risk perceptions and quit intentions via memory, but memory indices may differ in their predictive power.

Method

Participants and design

Participants were 866 adult smokers recruited using Amazon Mechanical Turk who participated online. Participants were at least 18 years old, had smoked at least 100 cigarettes in their lifetime, and currently smoked most days or every day. Eleven people were eliminated due to a programming error. Of those remaining, 725 (85%) completed some or all of the posttest measures.

Participants were randomly assigned to a 3 (Warning label (WL): High-emotion, Low-emotion, Text-only) × 2 (Posttest timing: Immediate vs. 6-week delay ×2 (Risk information order: Percentage first or Frequency first) design. Immediate vs. 6-week delay was assigned at a rate of 3:5 so that the two posttest conditions would have roughly equal numbers, after inevitably lower retention in the delay group. See Figure S1 for participant flow diagram. Participants received risk information in Percentage (e.g., 9%) or Frequency (e.g., 9 out of 100) formats, with participants randomized to receive either Percentage first or Frequency first. This manipulation of risk information format did not have the expected effects and is not discussed further in text (see SI1). The research was approved by the Ohio State University IRB (protocol 2010B0341).

Procedure

Before viewing the WLs, participants completed a measure of health literacy and answered questions about demographics, current tobacco use, quit intentions, and risk perceptions. In the same session, participants were exposed to eight WLs in random order four times each. The first time, they simply viewed all eight labels. In the second exposure, they rated their emotional reactions to each label, in the third exposure, they rated the credibility of each label, and in the fourth exposure, they rated the personal relevance of each label.

Finally, after all four exposures, participants either proceeded immediately to the posttest or completed it after a six-week delay as in prior research (Peters et al., 2019; Steidl et al., 2011). The posttest included free recall of smoking-related risk outcomes, cued recall and recognition of the specific risk probabilities for each risk outcome, quit intentions, risk perceptions, and objective numeracy. Both pretest and posttest also included other measures not used in the present research (see Table S1).

Materials

Warning labels. In three between-participants conditions, we varied the intended emotional reaction elicited by the eight labels. In the Text-only condition, participants viewed only text warnings (see below) and the eight risk-outcome and risk-probability statements. In the Low-emotion and High-emotion conditions, text warnings and risk information were paired with relevant images pretested to elicit little or a great deal of emotion, respectively. Low-emotion images included symbolic images (e.g., a cartoon gravestone) or mild outcomes (e.g., a man coughing); High-emotion images depicted severe outcomes such as damaged organs (e.g., a necrotic lung) or surgery (Peters et al., 2019; Evans et al., 2017, 2018). A sample of undergraduate students rated the Low-emotion images as less intense ($\text{Mean}_{\text{intensity}}$ = 1.79, $\text{Range}_{\text{intensity}}$ = 1.29 to 2.15) than the High-emotion images ($\text{Mean}_{\text{intensity}}$ = 4.50, $\text{Range}_{\text{intensity}}$ = 3.69 to 5.00; see SI 2 for additional pretesting details).

To create eight warnings, four congressionally-mandated text warnings were paired with risk-probability information for smokers and nonsmokers for two risk outcomes from existing risk evidence (Jha et al., 2013; Woloshin, Schwartz, & Welch, 2008). For example, the warning "smoking can kill you" was paired with "21.0% of smokers will die in the next 10 years, compared to 11.8% of non-smokers" and "75.4% of smokers will die before the age of 85, compared to 53.7% of non-smokers" or "21.0% of smokers will die in the next 10 years, compared to 11.8% of non-smokers." Thus, participants received each text warning twice and each risk outcome once per exposure (Table S2).

Measures

Emotional reactions. During the second exposure to the labels, participants used self-assessment manikins to indicate the arousal level (1 = *calm*, 5 = *excited*) of their emotional reactions to each label (Bradley & Lang, 1994). The emotional reaction score was the average arousal reaction across all eight labels, consistent with prior research (e.g., Peters et al., 2019).

Objective numeracy (ONS). At the end of the posttest, participants completed a seven-item Rasch-based numeracy measure (Weller et al., 2013). The items were modified slightly so that correct responses could not be found easily online as is the case for the original items (e.g., "On a bingo game show, the chance of winning an LED TV is 1 in 1,000. What percent of people on that bingo game show would win the LED TV on average?"). Numeracy scores were the number of correct answers out of 7.

Health literacy. Health literacy was assessed with Passage B from the Short Test of

Functional Health Literacy (Baker, Williams, Parker, Gazmararian, & Nurss, 1999). Passage B uses a modified cloze procedure that omits every fifth to seventh word of sentences written at a 10th grade reading level. Participants selected the best word from four options to complete the sentence (e.g., "I agree to give correct information to ______ if I can receive Medicaid."). Because total correct responses were highly skewed (roughly half of participants answered all 20 items correct), we rescored health literacy as a binary variable: 1=all items correct and 0=one or more items incorrect (see e.g., Lazard, Byron, Peters, & Brewer, 2019 for a similar approach). Controlling for health literacy as a continuous variable produced substantially similar results, but the structural equation model (see below) would not converge.

Risk perceptions. Participants answered five questions about perceived smoking risks (e.g., "Compared to the average nonsmoker your age, gender, and race, how would you rate your chances of dying from lung cancer" −3 = *Much lower*, +3 = *Much higher*; "If I don't stop smoking, I would feel very vulnerable to dying at a younger age because of smoking" −3= *strongly disagree*, +3 = *strongly agree*, Dillard, Ferrer, Ubel, & Fagerlin, 2012; Peters et al., 2019). Items were rescaled to a 1-7 scale so that higher numbers indicate greater risk perceptions. The average of the five items comprised the final risk-perception index.

Quit intentions. Quit intentions included a contemplation ladder (Biener & Abrams, 1991) about the participant's current action with respect to quitting (0 = *no thoughts of quitting*, 10 = *taking action to quit*) and two additional items based on past research (Romer, Peters, Strasser, & Langleben, 2013): "How likely do you think it is that you will try to quit smoking within the next 30 days [next year]?" ("−3 =*Very unlikely*," "+3 =*Very likely*"; rescaled to 1-7). Because they were measured on different scales, the quit intentions ladder was rescored to be on the same 7-point scale as the other two items by multiplying using the following formula:

y_2 = (range y_2)*(y_1)/(range y_1)) + miny$_2$,

where y_2 is the new rescaled value, range of y_2 is the range of the other items = 6, y_1 is the original value, range of y_1 is the range of the original scale = 10, and miny2 is the minimum value the rescaled variable can take = 1 (e.g., 0 became 1, 1 became 1.6, 10 became 7, etc.). The three items were averaged into an index in which higher numbers indicated greater intentions.

Risk Memory

Nonnumeric risk outcome recall. At posttest, participants were asked to recall as many smoking risk outcomes as possible. Responses were coded by two independent coders for mentions of 40 risk categories. The two coders agreed on 99.1% of trials and resolved disagreements through discussion. Seven risk outcomes were mentioned in the risk information accompanying the labels: death (mentioned in two labels), COPD, lung disease, heart disease, stroke, lung cancer, and cancer (Table S2). Scores on this measure were the proportion of the seven possible risk outcomes mentioned. Participants received a maximum of one point for each risk outcome recalled, even if they mentioned multiple types of the risk outcome. For example, a participant who mentioned mouth cancer and throat cancer was scored as correct (=1) for cancer.

Numeric risk probability recall. To measure posttest smoker and nonsmoker risk-probability recall, participants were asked eight fill-in-the-blank questions about the probabilities of each risk outcome separately for smokers and nonsmokers (e.g., "What percent of smokers [How many smokers] will die of heart disease?", "What percent of non-smokers [How many non-smokers] will die of heart disease?"). Very few participants recalled the exact digits from the labels (5% of answers were recalled exactly). Thus, answers were coded as correct (=1) if their answer was +/-5% from the correct answer (e.g., for smoker risk of dying before age 85, because the probability provided to participants was 75.4%, answers falling between 70.4% and

80.4% were coded as correct=1). Answers outside this range were considered incorrect and scored as 0. Risk probability recall scores were the proportion correct (out of eight); smoker and nonsmoker probability recall were scored and treated as separate variables.[1]

Numeric risk probability recognition. Participants were given a multiple-choice test of the probabilities for each risk outcome at posttest. There was one question for each label (i.e., eight total) that included smoker and nonsmoker risk (e.g., "______% of smokers will die before age 85; ______% of nonsmokers will die before age 85"). Each question had four decoy options. Risk probability recognition was the proportion of correct answers (=1) out of eight.

For all items with a correct response (memory, numeracy), missing responses were coded as incorrect.

Data analysis strategy

To test memory-consolidation H1 (that participants given High-emotion WLs would recall less immediately but have better retention over time (vs. WLs that elicited less emotion), generalized estimating equation (GEE) analyses with an identity link were carried out using SPSS v.25 (IBM, 2017). Specifically, we tested an interaction of WL (High-emotion, Low-emotion, and Text-only) and delay (Immediate or Six-week delay) conditions on memory for all risk memory indices. The four memory indices (i.e., risk outcome recall, smoker risk probability recall, nonsmoker risk probability recall, and risk probability recognition) were treated as repeated measures in this model. Non-significant effects were removed one at a time, with the exception of covariates (i.e., age, gender, race; see SI 3).

---

[1] We also tried a more liberal coding scheme that allowed responses +/-10% of the correct answer to be coded as correct. This scheme increased the average score on the recall measures from 33% to 48% for smoker risks and 54% to 72% for nonsmoker risks. The pattern of results is substantially the same with both rules. We chose to present results using the more conservative rule.

To test H2 (that greater numeracy would be associated with superior all-risk memory, controlling for education and health literacy) and H3 (that more accurate all-risk memory would be associated with greater risk perceptions and quit intentions and that memory would mediate the relation of numeracy with risk perceptions and quit intentions), we conducted a single structural equation model (SEM) using MPlus (Muthén & Muthén, 2012), with maximum likelihood estimation and robust (Huber-White) standard errors. We used model fit indices with cut-off values, [RMSEA]<=.05 with 90% Confidence Intervals of the RMSEA=.00–.08, Comparative Fit Index [CFI]>=.95 and Standardised Root Mean Square Residual [SRMR] <= .08 (Hu & Bentler, 1999). Indirect effects from numeracy to risk perceptions and quit intentions were estimated using bootstrapping with 5,000 resamples to evaluate significance. Our starting model included paths from numeracy (plus education and health literacy) to a latent memory variable constructed from all four memory indices, risk perceptions, and quit intentions. Paths were also included from the latent memory variable to risk perceptions and quit intentions separately. As in prior research (e.g., Peters et al., 2019), two warning-label condition contrasts were created. One contrasted pictorial WL conditions and Text-only condition (High-emotion=+1, Low-emotion=+1, Text-only = −2); the other contrasted High-emotion (+1) and Low-emotion (−1) labels (Text-only=0). Paths were included from the two warning-label condition contrasts (i.e., image vs. text and high-vs. Low-emotion) to emotional reaction and from emotional reaction to memory, risk perceptions, and quit intentions to examine whether numeracy' relations were robust to other known predictors of our outcome variables (e.g., Evans et al., 2015, 2017). Delay was included as a covariate for memory, as memory declines over time, and demographic variables (i.e., age, gender, race, baseline smoking heaviness) were included as covariates for outcomes. Nonsignificant paths were dropped one at a time based on the highest p-value, with the exception of demographic covariates. For paths involving numeracy, when numeracy was dropped, health literacy and education were also dropped, except when education was included as a demographic covariate.

We tested H3A (exploring predictive differences between the four memory indices for risk perceptions and quit intentions) using SEM. The only difference between this and the prior model was memory was not treated as a latent variable here. Instead, paths were included from objective numeracy (plus education and health literacy) to each separate memory index and from each memory index to risk perceptions and quit intentions.

Results

Of the 850 participants who completed the baseline measures, 18 were eliminated for viewing two or more initial WLs for less than 500ms, which was deemed insufficient for adequately processing the labels' number of words. Additionally, four participants who completed posttest measures were removed because they had 100% recall for 14-16 of the 16 risk-probability recall items (14 *SD*s above the mean), suggesting they wrote information down rather than remembering it. Our analyses focused on the 696 remaining participants.

Demographics

Participants (N=696) averaged 35.6 (SD=10.8) years of age and were 46.3% female, 80.3% white, 5.2% black, 6.6% Hispanic and 8.0% other; 82.6% had more than a high school education. They smoked an average of 11.6 (SD=7.9) cigarettes per day. We examined variables to determine whether they predicted return to the delayed posttest (all participants returned to the immediate posttest). Older age (and not gender, education, race, smoking heaviness, health literacy, or WL condition) predicted greater return (b=0.05(0.01), $p < .001$). Demographics and descriptive statistics appear in Table 1[2]; correlations are in Table S3.

Table 1. Descriptive statistics and demographic variables. Memory, risk perceptions and quit intentions were measured either immediately after the final WL exposure or after 6 weeks. Means (standard deviations) are provided for the overall sample and for the three WL conditions. Where relevant, Cronbach's α are provided.

			Mean (SD)			
	Cronbach's α	Range	Overall	Text-only	Low-emotion	High-emotion
Numeracy	.66	0-7	3.04 (1.68)	2.95 (1.74)	3.20 (1.60)	2.97 (1.69)
Health literacy binary	.75	0-1	0.53 (0.50)	0.54 (0.50)	0.56 (0.50)	0.50 (0.50)
Emotional reaction	.92	1-5	2.99 (0.88)	2.81 (0.87)	3.00 (0.76)	3.19 (0.94)
Risk outcome recall	.25	0-1	0.52 (0.19)	0.53 (0.18)	0.53 (0.20)	0.50 (0.19)
Smoker risk probability recall	.70	0-1	0.33 (0.27)	0.46 (0.26)	0.48 (0.26)	0.42 (0.26)
Nonsmoker risk probability recall	.67	0-1	0.54 (0.26)	0.33 (0.27)	0.34 (0.27)	0.30 (0.25)
Risk probability recognition	.62	0-1	0.45 (0.26)	0.56 (0.25)	0.56 (0.27)	0.51 (0.28)
Risk perceptions	.93	1-7	5.67 (1.15)	5.63 (1.21)	5.67 (1.09)	5.7 (1.16)
Quit intentions	.76	1-7	4.87 (1.62)	4.70 (1.74)	4.94 (1.58)	4.99 (1.51)
Smoking heaviness	--	0 – 45	11.62 (7.91)	12.67 (8.29)	11.09 (7.49)	10.98 (7.8)
Age	--	18 – 72	35.55(10.82)	36.8 (11.55)	35.67(10.99)	34.03(9.57)
Race (Non-Hispanic white)	--		80.3%	78.5%	83.7%	78.8%
Gender (female)	--		46.3%	43.50%	46.3%	49.6%
Education (> high school)	--		82.6%	80.2%	83.3%	84.7%

Test of H1 (that participants given High-emotion WLs would remember less immediately but have better retention over time (vs. Low-emotion participants)

[2] We conducted three 3 (WL condition) × 2 (Delay condition) ANCOVAs on emotional reactions, risk perceptions, and quit intentions (estimated means in Table S4). Emotional reactions differed by condition, $F(2,688)=10.81$, $p<.001$. Pairwise comparisons demonstrated that emotional reactions were highest in the high-emotion condition (*M(se)* = 3.19 (0.06), which were higher than low-emotion (*M(se)* = 3.00 (0.06), $p=.023$, which were higher than text-only (*M(se)* =2.82 (0.06), $p=.020$; high-emotion was higher than text only $p<.001$. Risk perceptions were similar across conditions. An interaction emerged of delay and warning condition on quit intentions, $F(2,678)=3.91$, $p=.021$ In the immediate posttest, an effect emerged of WL condition on quit intentions, $F(2,678)=3.07$, $p=.047$, such that quit intentions were significantly higher in the high-emotion condition (*M(se)*=5.13 (0.16)) than either of the other two conditions (low: *M(se)* =4.61 (0.16), $p=.025$; text-only: *M(se)*=4.67 (0.16), $p=.042$) which did not differ from one another ($p=.80$). Estimated marginal means in Table S4.

We used GEE to test for the expected interaction of delay and warning condition across the four memory indices (see Figure 1 and Table S4). First, average risk memory was lower in the delayed (*M(se)*=0.36 (0.01)) vs. immediate condition (*M(se)*=0.59(0.01), *Wald* $\chi^2(1)=324.90$, $p < .001$). It also was lower in the High-emotion condition (*M(se)*=0.45(0.01)) than in the Low-emotion condition (*M(se)*=0.49 (0.01)) or Text-only condition (*M(se)*=0.49 (0.01) (warning condition: Wald $\chi^2(2) = 9.88$, p =.007; condition comparisons: high:low, (b(se) = −0.08(0.03), p = .002; high:Text-only, b(se)= −0.06 (0.03), p =.021; low:Text-only, b(se) = 0.02 (0.02), p =.335.

H1 was partially supported (WL condition × delay condition: Wald $\chi^2(2)=5.88$, *p* =.053). Smokers in the immediate condition remembered less if they were in the High-emotion than other conditions (high:low, (b(se) = −0.09 (0.03), *p* =.001; high:Text-only: *b(se)*= −0.06 (0.02), *p* = .023; low:Text-only: *b(se)* =0.03 (0.03), *p* =.276). In addition, the effect of delay was weaker (i.e., memory was better supported) in the High-emotion condition than the Low-emotion or Text-only conditions (Immediate:Delay contrast in high, *b(se)* = −0.18(0.02), $p < .001$; low, *b(se)* = −0.26(0.02), $p < .001$; and Text-only, *b(se)* = −0.23(0.02), $p < .001$). Planned pairwise contrasts revealed that delay effects were different between the high:low conditions, (*b(se)* = −0.08 (0.03), *p*=.015), but not high:Text-only, *b(se)* = −0.04 (0.03), *p*=.206 or low:Text-only, *b(se)* = 0.04 (0.03), *p* =.192. However, the effect of WL condition was *ns* at six weeks (Wald $\chi^2(2)=1.93$, *p*=.381), indicating a possible floor effect. White participants had better memory, *b(se)* = 0.05 (0.02), *p* = .001; age and gender did not predict memory (*p*'s >.10).[3]

[3] Including memory index (risk-outcome recall, smoker risk recall, nonsmoker risk recall, risk recognition) as a factor did not reduce the WL × delay interaction. A nonsignificant three-way interaction of memory index × delay × warning condition, Wald $\chi^2(6) = 4.58$, *p* =.60, indicated that delay and WL condition effects were similar across memory indices.

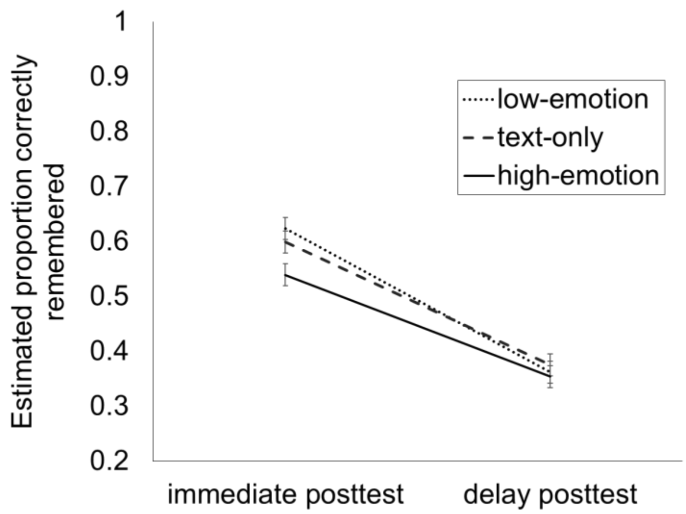

Figure 1. Estimated marginal mean proportion of all risk memory correctly remembered by WL and delay condition, controlling for race, age, and gender. Error bars indicate ± 1 *SE*.

Tests of H2 (that greater numeracy would be associated with superior all-risk memory, controlling for education and health literacy) and H3 (that more accurate all-risk memory would be associated with greater risk perceptions and quit intentions and that memory would mediate the relation of numeracy with risk perceptions and quit intentions)

We used a single SEM to test H2 and H3. Risk memory was estimated as a latent variable, and we modeled the effects of emotional reactions and numeracy on memory, risk perceptions, and quit intentions (see Figure S2 and Table S5 for starting model results; covariates included demographics, health literacy, and delay). The final model fit reasonably well (χ^2[65] = 257.55, p < .001; RMSEA = .066 [CI90: .058 to .075]; CFI = .88; SRMR = .04, see Figure 2 for main results and Table S6 for all paths, including covariate effects, factor loadings for latent memory variable, and indirect effects).

Participants in pictorial WL conditions reported greater emotional reaction than Text-only participants ($b(se)$ =0.09(0.02), p < .001). High-emotion participants reported greater emotional reaction to their warnings than Low-emotion participants ($b(se)$ =0.09(0.04), p = .031). Consistent with H2, more numerate individuals had more accurate all-risk memory, controlling for health literacy, education, and other socio-demographic and smoking variables ($b(se)$ = 0.01(0.002), p < .001; see Table 6 for effects of covariates). Health literacy was also associated with better memory ($b(se)$ =0.02(0.01), p = .009). Delay reduced memory ($b(se)$ = −0.10(.01), p < .001). The path from emotional reaction to memory was removed from the model.

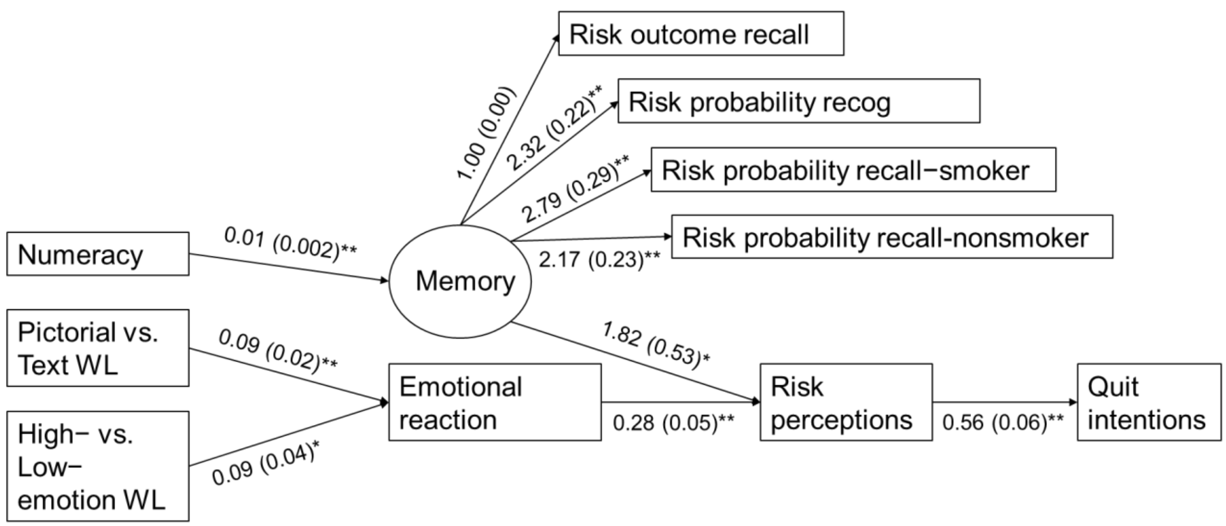

Figure 2. Final SEM predicting antecedents and consequences of risk memory. Path coefficients are unstandardized with standard errors indicated, *p < .05, **p < .01. See Table S6 for covariate and indirect effects.

Consistent with H3 (that memory effects would be consequential), accurate all-risk memory also predicted greater risk perceptions ($b(se)$ =1.82(0.53), p = .001). In addition, greater emotional reactions predicted greater risk perceptions ($b(se)$ =0.28(0.05), p < .001); emotion's effect was stronger (standardized $\beta(se)$ = .21(.04)) than that of memory ($\beta(se)$ = .13(.04)). The direct path from numeracy to risk perceptions was trimmed from the model. Quit intentions then were predicted by risk perceptions (*b(se)* =0.56(0.06), $p < .001$); direct paths to quit intentions from memory indices, emotional reaction, and numeracy were nonsignificant and trimmed from the model. H3 was supported: Numeracy indirectly predicted greater risk perceptions through memory (*indirect effect [IE]*: =0.03, *95%CI*: 0.01, 0.05). It also predicted greater quit intentions through memory then risk perceptions (*IE* =0.02, *95%CI*: 0.01, 0.03).

Exploratory analyses of H3A tested the same starting model but with memory indices allowed to covary rather than modeled as a single latent variable (Table S7). The results for paths included in the prior SEM were largely the same. The final model fit well: N=673, (χ^2[25] = 42.74, p =.015; RMSEA = .03 [CI90: .01 to .05]; CFI = .99; SRMR = .02; BIC = 5,238.53; Figure S3, Table S8). Numeracy predicted accurate memory for numeric risk-probability indices more strongly than it did for the nonnumeric risk-outcome index[4]. The effects of the memory indices on risk perceptions also differed. Specifically, greater risk-outcome recall (*b(se)* = 1.68(0.25), $p < .001$) and better risk-probability recognition (*b(se)* = 0.74(0.23), $p = .001$) both led to greater risk perceptions whereas better recall of smokers' risk probabilities was associated with *lower* risk perceptions (*b(se)* = −0.65(0.20), $p = .001$) and better recall of nonsmokers' risk probabilities was unrelated to risk perceptions. The total indirect effect of numeracy on risk perceptions was positive (total *IE*: =0.05, 95%CI: 0.03, 0.07) through risk-outcome recall (*IE*

---

[4] We also tested for an interaction between numeracy and memory index (risk-outcome recall [nonnumeric], risk-probability recognition [numeric], smoker risk-probability recall [numeric], and nonsmoker risk-probability recall [numeric]) using GEE analysis. The effect of numeracy differed across the memory indices, Wald $\chi^2(3)$=30.34, p <.001. In particular and as we expected, the effect of numeracy was stronger for each of the numeric risk probability indices than it was for the non-numeric risk outcome recall (risk probability recognition vs. non-numeric risk outcome recall, interaction *b(se)*= 0.03 (0.01), p <.001; smoker risk probability recall vs. non-numeric risk outcome recall, interaction *b(se)* = 0.01 (0.01), p =.042; nonsmoker risk probability recall vs. non-numeric risk outcome recall, interaction *b(se)* = 0.02 (0.01), p <.001). Additional details are provided in SI 4.

=0.03, 95%CI: 0.02, 0.05) and risk-probability recognition (*IE*=0.03, 95%CI: 0.01, 0.06), and negative through recall of risk probabilities for smokers (*IE* =−0.02, 95%CI: −0.03, −0.01).

Discussion

Pictorial WLs can improve risk knowledge (operationalized as memory in this paper), but not all pictorial WLs do so or do so immediately (Evans et al., 2015; Noar et al., 2016, 2017). Consistent with memory-consolidation findings (Peters et al., 2019), High-emotion warnings were remembered less well than Low-emotion warnings immediately, and memory for High-emotion warnings declined less over a six-week delay, thus partially supporting H1. However, unlike H1 and prior research, memory did not differ by condition at delay, a possible floor effect inconsistent with previous research using similar warnings but without added numeric information (e.g., Evans et al., 2015; 2017; Peters et al., 2019). The present participants had greater emotional reaction to High-emotion than Low-emotion WLs, but had the weakest reaction to Text-only WLs (vs. prior studies in which Low-emotion WLs elicited the lowest reaction). We may have overwhelmed participants with information. The current research exposed participants to 16 risk probabilities in addition to typical warning text, and it appeared that the frequency risk information was particularly difficult (SI 1). Additionally, in this heavily numeric study, reported emotional reactions may have been due to the warnings (as intended) and to the presence of numbers, particularly among the less numerate who report more negative emotions to them (Peters & Bjälkebring, 2015). In the present research, the less numerate, in fact, reported greater negative emotional reactions (Table S3). Across these studies, we conclude that emotional reaction drives immediate memory and memory changes over time, and that Low-emotion WLs can be more or less emotional than Text-only WLs.

Consistent with H2, greater numeracy was associated with better memory for risk information overall, controlling for health literacy and education. Health literacy and education also predicted information retention, but to a lesser extent. Future research should use a more difficult measure of health literacy (or a less educated sample) and examine whether it reduces numeracy's effect and especially for non-numeric memory. Numeracy's effect was larger for numeric than non-numeric risk memory indices but it was positive for all four memory indices, suggesting that those higher in numeracy also paid more attention to risk outcomes compared to those lower in numeracy (e.g., Fleig et al.2017; Jasper et al., 2017).

Consistent with H3 and prior research (Dillard et al., 2006; Peters et al., 2019), greater memory for presented risk information was associated with greater risk perceptions and, in turn, intentions to quit smoking. Indirect effects of numeracy suggest that more numerate individuals retained more information from WLs, which led them to perceive greater smoking risk, and, in turn, to have higher quit intentions. Thus, past research linking quitting to numeracy (Martin et al., 2012) and education (e.g., Siahpush et al., 2006) may be driven in part by more (vs. less) numerate smokers learning more from health warnings (Lillard, 2017).

Exploratory analyses testing H3A (that numeracy-mediated effects on risk perceptions and quit intentions may differ across memory indices) revealed potential differences in the consequences of memory indices. This model explained the data better than the model that treated risk memory as a latent variable, suggesting that separating memory indices added explanatory power. First, greater recall of risk outcomes predicted greater risk perceptions and, in turn, greater intentions to quit smoking (Dillard et al., 2006; Peters et al., 2019). Second, better recognition of risk probabilities also predicted greater risk perceptions and, in turn, greater quit intentions. This finding is novel, as prior research has not examined memory for risk probabilities. More accurate recall of nonsmokers' risk probabilities did not predict risk perceptions. Smokers may perceive these risks as non-self-relevant, despite them providing a benchmark to gauge smoking's increased risk. Finally and paradoxically, more accurate smoker risk probability recall predicted *lower* risk perceptions. Thus, numeracy both increased and increased risk perceptions via different memory indices. Because those higher in numeracy likely learn risk information better and make more numeric comparisons (Peters & Bjälkebring, 2015), they may be more likely to learn new risk information but they may also be more likely to be surprised and relieved by how low some of the risks were. Indeed, people sometimes overestimate smoking risk probabilities when asked to make risk estimates (Kristiansen et al., 1983; Viscusi, 1990) and past research has demonstrated that when people are relieved when they encounter numeric risk information, they have lower risk perceptions (Fagerlin, Zikmund-Fisher, & Ubel, 2005). Future research should investigate reactions to numeric risks such as surprise or relief. Communicating risk information that is lower than participants expect may reduce risk perceptions, at least short-term. However, once people are aware of the true probabilities, accuracy should no longer be associated with lower risk perceptions. Despite accurate smoker risk probability recall predicting lower risk perceptions, total positive indirect effects exist for numeracy on risk perceptions and quit intentions through risk memory, suggesting overall benefits to improving knowledge.

Although our study design included multiple exposures, it did not replicate naturalistic exposure to WLs on cigarette packages, which typically involves viewing single labels repeatedly while using the same cigarette pack. Our participants instead viewed multiple WLs repeatedly in a single session, and were asked questions immediately or after a six-week delay during which they did not see any of our WLs. We also did not assess actual quitting-related behavior. Finally, our sample was not nationally representative, although we did control for demographic variables, like education, that could have influenced the results.

Because smokers tend to have lower numeracy and education than nonsmokers (e.g., Cutler & Lleras-Muney, 2010; Martin et al., 2012), it may be tempting to conclude that risk messages should be nonnumeric because those lower in numeracy learned risk probabilities less well and more accurate memory of smoker risk probabilities led to lower numeric risk perceptions. More research is needed, however, to determine when, for how long, and for whom risk information will reduce risk perceptions (see SI 5; Table S10). The better long-term approach, however, is to teach accurate risk probabilities to increase informed choice while engaging smokers with information that deepens their currently superficial knowledge about the health risks of their habit. In particular, communicators should consider how best to support accurate risk understanding among the less numerate (e.g., presenting only the most important and relevant numeric information in a clear, easily evaluable format, see Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007; Zikmund-Fisher, Fagerlin, & Ubel, 2007; manipulations that reduce threat evoked by health warnings (e.g., self-affirmations), see Peters et al., 2017).

References

Ayanian, J. Z., & Cleary, P. D. (1999). Perceived risks of heart disease and cancer among cigarette smokers. *JAMA, 281*(11), 1019-1021. doi: 10.1001/jama.281.11.1019

Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Education and Counseling, 38*(1), 33-42. doi: https://doi.org/10.1016/S0738-3991(98)00116-5

Biener, L., & Abrams, D. B. (1991). The Contemplation Ladder: Validation of a measure of readiness to consider smoking cessation. *Health Psychology, 10*(5), 360-365. doi: 10.1037/0278-6133.10.5.360

Borland, R., Yong, H.-H., Balmford, J., Fong, G. T., Zanna, M. P., & Hastings, G. (2009). Do risk-minimizing beliefs about smoking inhibit quitting? Findings from the International Tobacco Control (ITC) Four-Country Survey. *Preventive Medicine*, *49*, 219–223.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi: https://doi.org/10.1016/0005-7916(94)90063-9

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences, 21*(7), 294-299.

Cantrell, J., Vallone, D. M., Thrasher, J. F., Nagler, R. H., Feirman, S. P., Muenz, L. R., . . . Viswanath, K. (2013). Impact of tobacco-related health warning labels across socioeconomic, race and ethnic groups: Results from a randomized web-based experiment. *PLoS ONE, 8*(1), e52206. doi: 10.1371/journal.pone.0052206

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics, 29*(1), 1-28. doi:

https://doi.org/10.1016/j.jhealeco.2009.10.003

de Boer, M. R., Waterlander, W. E., Kuijper, L. D., Steenhuis, I. H., & Twisk, J. W. (2015). Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 4. doi: 10.1186/s12966-015-0162-z

Dieckmann, N. F., Slovic, P., Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis, 29*(10), 1473-1488. doi: 10.1111/j.1539-6924.2009.01279.x

Dillard, A. J., Ferrer, R. A., Ubel, P. A., & Fagerlin, A. (2012). Risk perception measures' associations with behavior intentions, affect, and cognition following colon cancer screening messages. *Health Psychology, 31*(1), 106-113. doi: 10.1037/a0024787

Dillard, A. J., McCaul, K. D., & Klein, W. M. P. (2006). Unrealistic optimism in smokers: Implications for smoking myth endorsement and self-protective motivation. *Journal of Health Communication, 11*, 93-102. doi: 10.1080/10810730600637343

Emery, L. F., Romer, D., Sheerin, K. M., Jamieson, K. H., & Peters, E. (2014). Affective and cognitive mediators of the impact of cigarette warning labels. *Nicotine & Tobacco Research, 16*(3), 263-269.

Evans, A. T., Peters, E., Shoben, A. B., Meilleur, L. R., Klein, E. G., Tompkins, M. K., & Tusler, M. (2018). Cigarette graphic warning labels increase both risk perceptions and smoking myth endorsement. *Psychology & Health, 33*(2), 213-234.

Evans, A. T., Peters, E., Shoben, A., Meilleur, L., Klein, E. G., Tompkins, M. K., . . . Tusler, M. (2017). Cigarette graphic warning labels are not created equal: They can increase or decrease smokers' quit intentions relative to Text-only warnings. *Nicotine & Tobacco*

Research, 19(10), 1155-1162. doi: 10.1093/ntr/ntw389

Evans, A. T., Peters, E., Strasser, A. A., Emery, L. F., Sheerin, K. M., & Romer, D. (2015). Graphic warning labels elicit affective and thoughtful responses from smokers: Results of a randomized clinical trial. *PLoS ONE, 10*(12), 1-23. doi: 10.1371/journal.pone.0142879

Fagerlin, A., Ubel, P. A., Smith, D. M., & Zikmund-Fisher, B. J. (2007). Making numbers matter: present and future research in risk communication. *American Journal of Health Behavior, 31*, S47-56.

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2005). How making a risk estimate can change the feel of that risk: shifting attitudes toward breast cancer risk in a general public survey. *Patient Education and Counseling, 57*(3), 294-299. doi: https://doi.org/10.1016/j.pec.2004.08.007

Finucane, M. L., Slovic, P., Mertz, C. K., Flynn, J., & Satterfield, T. A. (2000). Gender, race, and perceived risk: The 'white male' effect. *Health, Risk & Society, 2*(2), 159-172. doi: 10.1080/713670162

Fleig, H., Meiser, T., Ettlin, F., & Rummel, J. (2017). Statistical numeracy as a moderator of (pseudo)contingency effects on decision behavior. *Acta Psychologica, 174*, 68-79. doi: https://doi.org/10.1016/j.actpsy.2017.01.002

Flynn, J., Slovic, P., & Mertz, C. K. (1994). Gender, Race, and Perception of Environmental Health Risks. *Risk Analysis, 14*(6), 1101-1108. doi: doi:10.1111/j.1539-6924.1994.tb00082.x

Galesic, M., & Garcia-Retamero, R. (2011). Communicating consequences of risky behaviors: Life expectancy versus risk of disease. *Patient Education and Counseling, 82*(1), 30-35. doi: 10.1016/j.pec.2010.02.008

Halpern, M. T. (1994). Effect of smoking characteristics on cognitive dissonance in current and former smokers. *Addictive behaviors*, *19*, 209-217.

Hammond, D. (2011). Health warning messages on tobacco products: a review. *Tobacco Control, 20*(5), 327-337.

Hartshorne, J. K., & Germine, L. T. (2015). When Does Cognitive Functioning Peak? The Asynchronous Rise and Fall of Different Cognitive Abilities Across the Life Span. *Psychological Science, 26*(4), 433–443. https://doi.org/10.1177/0956797614567339

Hoover, D. S., Wetter, D. W., Vidrine, D. J., Nguyen, N., Frank, S. G., Li, Y., . . . Vidrine, J. I. (2018). Enhancing smoking risk communications: The influence of health literacy and message content. *Annals of Behavioral Medicine, 52*(3), 204-215. doi: 10.1093/abm/kax042

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Jasper, J. D., Bhattacharya, C., & Corser, R. (2017). Numeracy predicts more effortful and elaborative search strategies in a complex risky choice context: A process-tracing approach. *Journal of Behavioral Decision Making, 30*(2), 224-235. doi: 10.1002/bdm.1934

Jha, P., Ramasundarahettige, C., Landsman, V., Rostron B., Thun, M., Anderson, R. N., . . . Peto, R. (2013). 21st-Century hazards of smoking and benefits of cessation in the United States. *New England Journal of Medicine, 368*(4), 341-350. doi: 10.1056/NEJMsa1211128

Klein, E. G., Quisenberry, A. J., Shoben, A. B., Romer, D., & Peters, E. (2019). The influence of health numeracy and health warning label type on smoking myths and quit-related reactions. *Nicotine & Tobacco Research, 21(*7), 974-978 doi: 10.1093/ntr/nty207

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology, 65*, 190-193.

Kristiansen, C. M., Harding, C. M., & Eiser, J. R. (1983). Beliefs about the relationship between smoking and causes of death. *Basic and Applied Social Psychology, 4*(3), 253-261. doi: 10.1207/s15324834basp0403_4

Lazard, A. J., Byron, M. J., Peters, E., & Brewer, N. T. (2019). Communicating about chemicals in cigarette smoke: impact on knowledge and misunderstanding. *Tobacco Control*, tobaccocontrol-2018-054863. doi: 10.1136/tobaccocontrol-2018-054863

Lillard, D. R. (2017). Educational heterogeneity in the association between smoking cessation and health information. In K. Bolin, B. Lindgren, M. Grossman, D. Gyrd-Hansen, T. Iversen, R. Kaestner & J. L. Sindelar (Eds.), *Human Capital and Health Behavior (Advances in Health Economics and Health Services Research)* (Vol. 25, pp. 183-206): Emerald Publishing Limited.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Medical Decision Making, 30*(4), 464–473.

Martin, L. T., Haas, A., Schonlau, M., Derose, K. P., Rosenfeld, L., Rudd, R., & Buka, S. L. (2012). Which literacy skills are associated with smoking? *Journal of Epidemiology and Community Health (1979-), 66*(2), 189-192.

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework.

Perspectives on Psychological Science, 2(1), 33-52. doi: doi:10.1111/j.1745-6916.2007.00028.x

Muthen L.K & Muthen B.O.. *Mplus User's Guide (Sixth Edition)*. Los Angeles, CA: Muthen & Muthen; 2012.

Noar, S. M., Hall, M. G., Francis, D. B., Ribisl, K. M., Pepper, J. K., & Brewer, N. T. (2016). Pictorial cigarette pack warnings: a meta-analysis of experimental studies. *Tobacco Control, 25*(3), 341-354. doi: 10.1136/tobaccocontrol-2014-051978

Noar, S. M., Francis, D. B., Bridges, C., Sontag, J. M., Brewer, N. T., & Ribisl, K. M. (2017). Effects of strengthening cigarette pack warnings on attention and message processing: A systematic review. *Journalism & Mass Communication Quarterly, 94*(2), 416-442.

Nonnemaker, J., Farrelly, M.C., Kamyab, K., Busey, A., & Mann, N. (2010). *Experimental Study of Graphic Cigarette Warning Labels: Final Results Report Prepared for Center for Tobacco Products, Food and Drug Administration*. Contract No.: Contract No. HHSF–223–2009–10135G.

Peters, E., Baker, D. P., Dieckmann, N. F., Leon, J., & Collins, J. (2010). Explaining the effect of education on health: A field study in Ghana. *Psychological Science, 21*(10), 1369-1376. doi: 10.1177/0956797610381506

Peters, E., & Bjälkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology, 108*(5), 802-822. doi: 10.1037/pspp0000019

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review, 64*(2), 169-190. doi: 10.1177/10775587070640020301

Peters, E., Hart, P. S., Tusler, M., & Fraenkel, L. (2014). Numbers matter to informed patient choices: A randomized design across age and numeracy levels. *Medical Decision Making, 34*(4), 430-442. doi: 10.1177/0272989x13511705

Peters, E., Lipkus, I., & Diefenbach, M. A. (2006). The functions of affect in health communications and in the construction of health preferences. *Journal of Communication, 56*, S140-S162. doi: 10.1111/j.1460-2466.2006.00287.x

Peters, E., Shoots-Reinhard, B., Evans, A. T., Shoben, A., Klein, E., Tompkins, M. K., . . . Tusler, M. (2019). Pictorial warning labels and memory for cigarette health-risk information over time. *Annals of Behavioral Medicine, 53*(4), 358-371. doi: https://doi.org/10.1093/abm/kay050

Peters, E., Shoots-Reinhard, B., Tompkins, M. K., Schley, D., Meilleur, L., Sinayev, A., . . . Crocker, J. (2017). Improving numeracy through values affirmation enhances decision and STEM outcomes. *PLoS ONE, 12*(7), e0180674. doi: 10.1371/journal.pone.0180674

Peters, E., Vastfjall, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and Decision Making. *Psychological Science, 17*(5), 407-413.

Petrova, D. G., Pligt, J., & Garcia- Retamero, R. (2014). Feeling the Numbers: On the Interplay Between Risk, Affect, and Numerac. *Journal of Behavioral Decision Making, 27*(3), 191-199. doi: 10.1002/bdm.1803

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973. doi: 10.1037/a0017327

Romer, D., Peters, E., Strasser, A.A., & Langleben, D. (2013). Desire versus efficacy in smokers' paradoxical reactions to pictorial health warnings for cigarettes. *PLoS ONE, 8*(1):

e54937. doi:10.1371/journal.pone.0054937.

Schwartz, P. H., Perkins, S. M., Schmidt, K. K., Muriello, P. F., Althouse, S., & Rawl, S. M. (2017). Providing quantitative information and a nudge to undergo stool testing in a colorectal cancer screening decision aid: A randomized clinical trial. *Medical Decision Making, 37*(6), 688-702. doi: 10.1177/0272989x17698678

Siahpush, M., McNeill, A., Borland, R., & Fong, G. T. (2006). Socioeconomic variations in nicotine dependence, self-efficacy, and intention to quit across four countries: findings from the International Tobacco Control (ITC) Four Country Survey. *Tobacco Control, 15*(suppl 3), iii71-iii75. doi: 10.1136/tc.2004.008763

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior, 24*(3), 271-296.

Steidl, S., Razik, F., & Anderson, A. K. (2011). Emotion enhanced retention of cognitive skill learning. *Emotion, 11*(1), 12-19. doi: 10.1037/a0020288

U.S. Department of Health and Human Services. (2014). *The health consequences of smoking--50 years of progress: A report of the Surgeon General*. Atlanta, GA: U.S. Dept of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014.

Wang, T. W., Asman, K., Gentzke, A. S., Cullen, K. A., Holder-Hayes, E., Reyes-Guzman, C., . . . King, B. A. (2018). Tobacco Product Use Among Adults - United States, 2017. *MMWR. Morbidity And Mortality Weekly Report, 67*(44), 1225-1232. doi: 10.15585/mmwr.mm6744a2

Weinstein, N. D. (1998). Accuracy of smokers' risk perceptions. *Annals of Behavioral Medicine,* *20*(2), 135-140. doi: 10.1007/bf02884459

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making, 26*(2), 198-212.

Woloshin, S., Schwartz, L. M., & Welch, H. G. (2008). The risk of death by age, sex, and smoking status in the United States: Putting health risks in context. *Journal of the National Cancer Institute, 100*(12), 845-853. doi: 10.1093/jnci/djn124

Viscusi, W. K. (1990). Do smokers underestimate risks? *Journal of Political Economy, 98*(6), 1253.

Zhuang, Y. L., Gamst, A. C., Cummins, S. E., Wolfson, T., & Zhu, S. H. (2015). Comparison of smoking cessation between education groups: Findings from 2 US national surveys over 2 decades. *American Journal of Public Health, 105*(2), 373-379. doi: 10.2105/ajph.2014.302222.

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2007). Mortality versus survival graphs: Improving temporal consistency in perceptions of treatment effectiveness. *Patient Education and Counseling, 66*(1), 100-107.