Journal of Experimental Psychology: Applied
2016, Vol. 22, No. 2, 247–260


1076-898X/16/$12.00 http://dx.doi.org/10.1037/xap0000086

# To Screen or Not to Screen: What Factors Influence Complex Screening Decisions?

Dafina Petrova
University of Granada

Rocio Garcia-Retamero
University of Granada and Max Planck Institute for Human Development, Berlin, Germany

Andrés Catena
University of Granada

Joop van der Pligt
University of Amsterdam

Contrary to people's intuitions, many screenings can have both benefits (e.g., lives saved) and harms (e.g., unnecessary treatments). Statistical information is often provided to ensure informed decision making. However, few theoretical models have addressed the role of comprehension of such information in screening decisions. In an experiment, we studied how cognitive skills, emotions, and a priori beliefs about screening affect comprehension of the evidence of benefits and harms from screening and intentions to get screened. Young adults ($N$ = 347) received information about a disease for which a screening test was available and numerical information about the benefits and harms from screening. Results showed that comprehension and perceptions of benefits are central to decisions; however, lay perceptions of harms along the screening cascade require further study. Numeracy, science literacy, and emotions can promote informed decision making by facilitating comprehension of the evidence. At the same time emotions and beliefs resulting from persuasive campaigns can have strong effects on screening intentions beyond the available evidence. To apply to screening procedures where informed decision making is recommended, theoretical models of screening decisions need to include comprehension of benefits and harms, and account for how cognitive skills, emotions, and beliefs influence comprehension and decisions.

*Keywords:* screening, numeracy, science literacy, emotions, risk literacy

*Supplemental materials:* http://dx.doi.org/10.1037/xap0000086.supp



Developments in medicine and health services have made an increasing number of screening tests available to individuals of various ages. Screening now starts in the womb (e.g., prenatal screening for Down's syndrome) and continues until much later in life (e.g., cancer screening). It is a good guess that our children will live in a world abundant with genetic screening tests. Screenings are done on asymptomatic people to find those at increased risk of having a disease or disorder, with the purpose to prolong and/or increase quality of life (Grimes & Schulz, 2002). As such, to be considered effective, screenings need to show benefit in randomized controlled trials (e.g., they should reduce mortality in people who go through screening compared to those who do not).

To the surprise of many, not all screenings show enough benefit to be recommended by authorities. For example, screening for prostate cancer with the prostate-specific antigen (PSA) test is considered not to be life-saving on average and it is not recommended in a number of countries (Ilic, Neuberger, Djulbegovic, & Dahm, 2013; Moyer, 2012). Even more counterintuitively, some screenings can cause harms to individuals and by extension to economies. These harms have recently begun to receive more attention and have been categorized into four broader types: physical harms, psychological harms, financial strain, and opportunity costs (Harris et al., 2014). While people may be used to drug treatments causing adverse effects, they may not expect that preventive procedures can be harmful. For example, to the surprise of some women, screening for breast cancer with mammography results in many false positive tests and causes a proportion of women to undergo unnecessary cancer treatment, including mastectomy (i.e., the so called

Dafina Petrova, Mind, Brain, and Behavior Research Center, University of Granada; Rocio Garcia-Retamero, Mind, Brain, and Behavior Research Center, University of Granada and Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany; Andrés Catena, Mind, Brain, and Behavior Research Center, University of Granada; Joop van der Pligt, Department of Social Psychology, University of Amsterdam.

The current research was funded by the Ministerio de Economía y Competitividad (Spain; PSI2011-22954 and PSI2014-51842-R). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Unfortunately, our co-author Prof. Joop van der Pligt passed away while the manuscript was still in preparation. We would like to acknowledge his indispensable help in the preparation and execution of this research.

Correspondence concerning this article should be addressed to Dafina Petrova, Mind, Brain, and Behavior Research Center, University of Granada, Granada 18071, Spain. E-mail: dafinapetrova@ugr.es

overdiagnosis bias; Gøtzsche & Jørgensen, 2013; Waller, Douglas, Whitaker, & Wardle, 2013).

Where difficult trade-offs between potential benefits and harms from screening need to be considered, experts recommend policies that promote informed decision making (Esserman, Thompson, & Reid, 2013; Rimer, Briss, Zeller, Chan, & Woolf, 2004; Sheridan, Harris, & Woolf, 2004). This means that patients need to consider the relevant benefits, harms, risks, and limitations, and make a decision consistent with their preferences (Rimer et al., 2004). The information that should be communicated to patients frequently includes complex medical terms and probabilistic, numerical information about the evidence of benefits and harms. Comprehension of such information is potentially central to informed decision making.

However, there is no encompassing research framework that addresses the impact of various psychological factors on comprehension and decisions about screening when numerical information about both benefits and harms needs to be considered. For example, several health behavior models have been utilized in the context of cancer screening decisions—including the Health Belief Model (Rosenstock, 1974) and the Theory of Reasoned Action (Fishbein, 1980). These theories have identified important constructs that affect screening intentions and behavior. Some examples are perceived benefits and costs of screening, beliefs about screening, and perceived severity of the disease that screening might detect. However, these theoretical models did not consider the role of comprehension of health-relevant information. Other health behavior models like Social Cognitive Theory (Bandura, 2004) or the Information-Motivation-Behavioral Skills model (Fisher & Fisher, 2002) have addressed the role of health relevant knowledge in the prediction of health behavior. However, to the best of our knowledge, these models have been mostly based on and applied in the context of persuasion-based health promotion (i.e., encouraging a health behavior that is deemed desirable by experts). They have also not emphasized the role of numerical risk and benefit information that is common for decisions for which informed decision making is recommended. In contrast, Fuzzy Trace Theory (Reyna, 2008), a model applied to medical decision making, gives a central role to comprehension. Fuzzy Trace Theory posits that decision makers rely on the gist of information (i.e., its bottom-line meaning) as opposed to verbatim details (i.e., the precise numbers). Research in physicians recently showed that gist knowledge was related to physicians' perceptions of benefits and harms from cancer screening, emphasizing the importance of comprehension in this context (Elstad et al., in press). However, while Fuzzy Trace Theory offers a comprehensive memory-based account of comprehension and how it affects decisions (Reyna, 2012), it does not make predictions about factors that may influence decisions without affecting gist or verbatim comprehension of the benefits and risks.

Other research showed how improved comprehension of benefits and harms can affect cancer screening intentions: Comprehension was related to more desire to participate in decision making about screening and better decisions (e.g., no intention to participate in a cancer screening program that offered no benefits; Petrova, Garcia-Retamero, & Cokely, 2015). Nevertheless, the effect of comprehension was limited among individuals who perceived cancer as extremely severe. This result suggests that beyond comprehension, emotional reactions to the prospect of a diagnosis and beliefs about screening can affect decision making, including the way in which individuals approach information about screening outcomes (Peters, Lipkus, & Diefenbach, 2006; Slovic, Peters, Finucane, & MacGregor, 2005). The psychology of modern screening decisions may benefit from a model that gives a central role to comprehension of the complex (numerical) information involved, and considers possible antecedents (e.g., emotional involvement, cognitive skills, and attitudes) and consequences (e.g., screening intentions, behavior). The purpose of this research was twofold. Our first aim was investigating the influence of various factors—cognitive skills, emotions, and beliefs—on comprehension of screening statistics and screening intentions. Our second aim was to start building a model of factors that influence complex screening decisions.

## Cognitive Skills

The necessity to comprehend complex numerical information in a medical context suggests that both numeracy and science literacy play an essential role in decision making. Numeracy refers to the ability to understand and operate with numerical and probabilistic concepts (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Galesic & Garcia-Retamero, 2010; Lipkus, Samsa, & Rimer, 2001; Peters & Bjalkebring, 2015), a skill that might be essential when considering benefits and costs of screening participation. People with low numeracy overestimate risks and benefits of treatments and are less capable of adjusting their estimates from risk reduction information (Davids, Schapira, McAuliffe, & Nattinger, 2004; Garcia-Retamero & Galesic, 2010; Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010; Reyna, Nelson, Han, & Dieckmann, 2009; Schwartz, Woloshin, Black, & Welch, 1997). People with low numeracy also search for less health-relevant information and prefer to leave decision making to their physician (Cokely et al., 2012; Galesic & Garcia-Retamero, 2011; Garcia-Retamero & Galesic, 2012; Garcia-Retamero, Wicki, Cokely, & Hanson, 2014). People with high numeracy, on the other hand, deliberate longer and make superior decisions (Ghazal, Cokely, & Garcia-Retamero, 2014).

While the predictive power of numeracy is well-established for various medical decisions, a substantial amount of variance in comprehension remains unexplained. The nature of the information and statistics about screening outcomes suggests that science literacy can also contribute to comprehension (Laugksch, 2000; National Science Foundation, 2014). In particular, understanding of how science generates and assesses evidence can help individuals evaluate the evidence of benefits and harms of screening (National Science Foundation, 2014, pp. 7–23). For example, when judging the effectiveness of screening, many people may consider all individuals alive in the screening group as "saved by screening"—even if there is a similar number of individuals alive in the nonscreening group (Petrova et al., 2015). This result suggests that knowing the essence of experimental methods (e.g., a control group is required to establish the benefit of a treatment) can improve comprehension and adjust perceptions. Furthermore, good science literacy may encourage people to approach rather than avoid health-relevant numerical information. For example, people with adequate science literacy may have more experience with and greater liking for scientific information. This might increase deliberation (e.g., the amount of time that people spend reading relevant information) and facilitate comprehension. However, it is still unclear whether the influence of science literacy is independent from that of numeracy. To fill this gap, in this research we tested the following hypothesis:

> *Hypothesis 1 (H1):* Higher science literacy will be related to longer deliberation and better comprehension of benefits and harms of screening, and this effect will be independent from that of numeracy.

## Emotions

Emotions can also affect screening intentions. Screening decisions are often emotionally laden because of the potential serious consequences (e.g., diagnosis of cancer; Zikmund-Fisher, Fagerlin, & Ubel, 2010). In the current research we tested two competing hypotheses about the effect of emotions on comprehension and decisions about screening. When the consequences of a decision are perceived to be emotionally powerful (e.g., fear-inducing), decision makers tend to pay less attention to numerical, probabilistic information and more often rely on heuristic-like processes that neglect the likelihood of specific events (Pachur, Hertwig, & Wolkewitz, 2014; Petrova, van der Pligt, & Garcia-Retamero, 2014; Rottenstreich & Hsee, 2001). This suggests that individuals who are more worried about a certain disease, or perceive the disease as more severe and unpleasant, may pay less attention to screening statistics and the actual evidence of benefits or harms. Instead, they would use heuristic-like strategies like "prevention is always good." This implies that strong emotional reactions can be detrimental to comprehension and increase screening intentions, even when the evidence is not in favor of screening. Alternatively, people who are more worried about a certain disease may also be more motivated to understand the relevant evidence and make informed decisions. In other words, emotions can function as "spotlight" or "motivators": they can increase people's interest and motivation to process the relevant information (Peters et al., 2006), which can have a beneficial effect on comprehension and decisions.

> *Hypothesis 2a (H2a):* Stronger emotional reactions will be related to worse comprehension and stronger screening intentions.

> *Hypothesis 2b (H2b):* Stronger emotional reactions will be related to better comprehension and weaker screening intentions.

## Beliefs

Beliefs about screening can also be related to comprehension and screening intentions. Many people can have strong positive beliefs about screening (e.g., many people think that screening is always a good choice; Schwartz, Woloshin, Fowler, & Welch, 2004; Waller, Osborne, & Wardle, 2015). These beliefs could rightfully stem from the perceived value of early prevention for saving people's lives. These beliefs can be further enforced by exposure to persuasive campaigns encouraging regular screening. However, such campaigns rarely mention potential harms or specify the exact degree of benefit (e.g., the "pink ribbon" campaigns for breast cancer screening; Gigerenzer, 2014). As a result, people may be left with the impression that screening is useful by definition rather than a matter of choice based on a cost-benefit analysis. To illustrate, many United States and European adults believe that cancer screenings are almost always beneficial and often grossly overestimate their benefits (Gigerenzer, Mata, & Frank, 2009; Hoffman et al., 2010; Schwartz, Woloshin, Sox, Fischhoff, & Welch, 2000; Schwartz et al., 2004; Waller et al., 2015). Similarly, many people perceive that screening is an obligation to one's family and society and perceive foregoing screening as irresponsible behavior (Hersch et al., 2013; Schwartz et al., 2004; Waller et al., 2015).

In our research we also tested two alternative hypotheses about the effect of beliefs on comprehension and screening intentions. In particular, nonevidence driven, prior beliefs can bias the processing of new information. For instance, prior beliefs might lead people to process new information about harms from screening more shallowly and/or discount the new, inconsistent information altogether (Garcia-Retamero, Müller, Catena, & Maldonado, 2009; Kunda, 1990; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). This implies that stronger positive a priori screening beliefs will reduce comprehension and, in turn, will increase screening intentions, even when the evidence is not in favor of screening. Alternatively, prior beliefs might not affect comprehension but only the way that people weigh potential benefits and harms from screening (Garcia-Retamero et al., 2009).That is, people with stronger positive beliefs might show similar levels of comprehension to those with less strong positive beliefs; however, they might focus on the benefits and ignore the harms when they make decisions about screening.

> *Hypothesis 3a (H3a):* Stronger positive a priori beliefs about screening will be related to lower comprehension and stronger screening intentions.

> *Hypothesis 3b (H3b):* Stronger positive a priori beliefs about screening will be related to stronger screening intentions, regardless of comprehension.

## Theoretical Model

We tested a model based on predictions derived from the Health Belief Model (Rosenstock, 1974), theories of numerical abilities, risk comprehension, and decision making (Reyna et al., 2009), and previous models of comprehension and informed decision making about screening (Petrova et al., 2015). People might differ in the amount of time that they spend deliberating on the statistical information that they receive. Based on previous research linking decision latency to superior decisions (Ghazal et al., 2014), we expected that shorter deliberation time would predict worse comprehension of screening outcomes, which in turn would predict more perceived benefits and fewer perceived harms from screening (Petrova et al., 2015). More perceived benefits and fewer perceived harms might be related to stronger intentions to get screened (Rosenstock, 1974). We further expected that participants would report stronger emotional reactions (e.g., fear, worry) for a disease described as more severe than for a disease described as rather neutral. These emotional reactions might be related to comprehension either negatively (H2a) or positively (H2b). We further tested whether individual differences in numeracy and science literacy would affect perceptions and intentions directly and indirectly (via comprehension). In particular, we expected numeracy and science literacy to be related to longer deliberations times and better comprehension (H1).We also tested whether a priori screening beliefs were related to stronger screening intentions indirectly (via comprehension; H3a) or directly (H3b).

To test this model under different conditions of screening effectiveness we manipulated the screening statistics. In one condition we presented participants with screening statistics showing that screening was effective at reducing mortality from the disease; in another condition screening was not effective. Hence, we investigated screening intentions when a normative decision option existed (i.e., not to participate in screening that is not effective) and under more ecological conditions where the extent of benefit needs to be weighed against the extent of possible harm (i.e., there is no "correct" decision option). We expected that the screening statistics manipulation would affect perceived benefits and intentions but not perceived harms (as the latter were held constant); we also accounted for a possible effect of the statistics manipulation on deliberation time and comprehension.

In summary, we investigated whether cognitive skills, emotions, and beliefs affect screening intentions by influencing comprehension of relevant evidence (i.e., *indirect effect on intentions via comprehension*). If these factors can affect screening intentions by impeding or facilitating comprehension of benefits and harms, results can help design interventions that promote informed decision making. If these factors have a direct effect on screening intentions, results can help identify psychological processes influencing screening intentions beyond the statistical evidence.

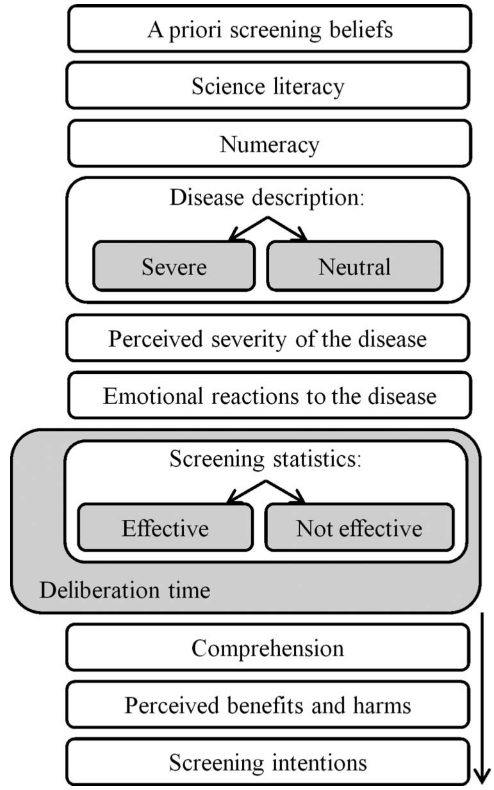


*Figure 1.* Design and procedure scheme of the experiment ($N$ = 347).

## Method

### Participants

Participants were 347 first-year psychology students at the University of Amsterdam in the Netherlands (66% women, age $M$ = 20, $SD$ = 2.60). They completed the study at the beginning of their first academic semester (October 2014, i.e., before receiving any substantial training in psychology and scientific methods) as part of the University of Amsterdam test sessions for new students. Data collection was approved by the ethics committee at the Department of Psychology of the University of Amsterdam.

### Stimuli and Design

Figure 1 shows a scheme of the design and the procedure of the experiment.

**Disease description.** Participants read a description of a hypothetical disease—Greene's disease—for which screening was available. Participants were told that the disease was discovered in Europe and affected around 5% of young adults. The disease was characterized by abnormal cells that were spread through the blood stream and could potentially cause death.

To investigate the effect of emotions, we manipulated the description of the disease. Half of the participants read a *severe* disease description, including extremely unpleasant symptoms of the disease and side effects of the treatment. The rest of the participants received a relatively more *neutral* description of similar length that did not include information about the unpleasant effects. The descriptions were designed to evoke strong and weak negative emotional reactions, respectively (see the Appendix for the exact text of both conditions).

**Screening statistics.** All participants were told that there was a simple blood test available that could screen for Greene's disease. Participants read a short explanation about the potential benefits and harms of screening. They were told that the blood test was developed to detect the disease at an early stage, when the treatment had more chances of being effective. Participants were explained that the test could also detect some abnormal cells that do not threaten the person's life. However, doctors could not differentiate between dangerous and nondangerous abnormal cells and people who test positive would get the treatment for Greene's disease even if they do not need it. Thus, some individuals would be treated unnecessarily.

All participants also read statistical information about the benefits and harms of screening for Greene's disease. The information was modeled after information available on government websites and decision aids about screenings with benefits and harms (see cdc.gov and harding-center.mpg.de). The statistics presented were fictitious but representative of the degree of harms and benefits from some cancer screenings (e.g., Gøtzsche & Jørgensen, 2013; Ilic, Neuberger, Djulbegovic, & Dahm, 2013). In particular, participants were told that scientists conducted a large-scale experiment to test the effectiveness of the blood test to save lives in 2,000 young people: One thousand were randomly assigned to participate in screening and another 1,000 were assigned not to participate in screening for 5 years.

***Benefits.*** The information about the benefits of screening was manipulated between-subjects. Half of participants read information showing that *screening was effective* (i.e., saved lives). In particular, these participants were told: "Of 1000 young people who participated in regular screening for 5 years, *3 people died* of Greene's disease. Of 1000 young people who did not participate in regular screening for 5 years, *5 people died* of Greene's disease." The rest of the participants read information showing that *screening was not effective*. These participants were told: "Of 1000 young people who participated in regular screening for 5 years, *5 people died* of Greene's disease. Of 1000 young people who did not participate in regular screening for 5 years, *5 people died* of Greene's disease."

***Harms.*** All participants received the same information about the degree of potential harm from screening. All participants were told: "From the 1,000 people who participated in screening, 6 people were treated with the laser treatment unnecessarily. From the 1,000 people who did not participate in screening, none were treated unnecessarily."

## Measures

We measured a priori screening beliefs, cognitive skills (i.e., science literacy and numeracy), perceived severity of the disease (manipulation check), emotional reactions to the disease, deliberation (i.e., time that participants spent reading the screening statistics), comprehension, perceived benefits and harms, and screening intentions.

**A priori screening beliefs.** On scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), participants indicated to what extent they agreed with six statements that reflected positive attitudes toward screening (e.g., "Screening always has more advantages than disadvantages" and "Foregoing screening is irresponsible;" see Appendix for a full description). One item was excluded for low item-to-total correlation. The final score was the average of the remaining five items (Cronbach's $\alpha = .80$) with $M = 4.46$, $SD = 1.14$.

**Science literacy.** It was assessed with three questions adapted from the US National Science Foundation survey (items 1–3, National Science Foundation, 2014) and two new items designed to complement the measure of the construct and increase reliability of the assessment (Items 4–5; see Appendix for a full description). The final score was a sum of the number of correct answers on the five questions (Cronbach's $\alpha = .63$) with a mean of 3.95 ($SD = 1.28$).

**Numeracy.** It was assessed with four items from the Berlin Numeracy Test (Cokely et al., 2012; see www.RiskLiteracy.org). We computed the number of correct answers for each participant. The final score had a mean of 1.30 ($SD = 1.04$).[1]

**Perceived severity of the disease.** Using scales from 1 = *not at all* to 7 = *most certainly*, participants rated Greene's disease on four dimensions (severe, serious, unpleasant, and horrible). We computed the average score on the four items (Cronbach's $\alpha = .83$). The final score had a mean of 5.25 ($SD = 1.11$).

**Emotional reactions to the disease.** On scales from 1 = *not at all* to 7 = *most certainly* participants indicated to what extent they would feel afraid, worried, angry, hopeful, and calm if they had Greene's disease. The positive items were reverse-scored. We computed the average of the five items (Cronbach's $\alpha = .64$) and the final score had a mean of 4.84 ($SD = .83$).

**Deliberation time.** We recorded how much time participants spent reading the screening statistics before they clicked "next" to see the comprehension questions. The comprehension questions were shown on a second page, and the screening statistics remained on the screen while participants answered them. However, participants were not informed that the statistics were going to remain on the screen while they answered the questions. Therefore, we used the time that participants spent reading the statistics on the first page as a proxy of deliberation. To correct for the positive skew typical for reaction time (RT) measures, the deliberation time measure was log-transformed for analysis.

**Comprehension.** Participants answered eight test-type questions that were adapted from previous research on screening with benefits and harms (see Appendix; see also Petrova et al., 2015). These items were designed to assess (a) comprehension of the idea that screening can have harms (Items 1–3) and (b) comprehension of the most essential screening statistics (Items 4–8).We computed the sum of correctly answered items. The items (Cronbach's $\alpha = .68$) showed good discriminability with a mean of 4.19 ($SD = 2.11$).

**Perceived benefit of screening.** On scales ranging from 1 (*not at all*) to 7 (*very much*) participants indicated (a) how effective screening with the blood test is, (b) how important, and (c) how beneficial it is to participate in screening for Greene's disease. The final score was an average of the three items (Cronbach's $\alpha = .84$) with a mean of 4.05 ($SD = 1.30$).

**Perceived harm from screening.** On scales ranging from 1 (*not at all*) to 7 (*very much*) participants indicated (a) how harmful and (b) how risky it is to participate in screening for Greene's disease. The final score was an average of the two items (Cronbach's $\alpha = .75$) with a mean of 3.79 ($SD = 1.13$).

**Screening intentions.** On scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*) participants indicated their agreement with the following statements: "I would regularly participate in screening for Greene's disease," "I would recommend to others to participate in screening for Greene's disease," "I would pay if necessary to participate in screening for Greene's disease." The final score was the average of the three items (Cronbach's $\alpha = .89$) with a mean of 3.51 ($SD = 1.29$).

In summary, we used a 2 (disease description: severe or neutral) × 2 (screening statistics: effective or not effective) between-subjects design. Individual differences in a priori screening beliefs and cognitive skills (i.e., science literacy and numeracy) were considered as independent variables in the analyses. The main dependent variable was screening intentions, while emotional reactions to the disease, deliberation time, comprehension, and perceived benefit and harm were considered as potential mediators. Perceived severity of the disease was considered as a manipulation check.

[1] We conducted exploratory and confirmatory factor analysis to verify that the science literacy and numeracy items reflect two distinct constructs. Exploratory factor analysis returned a two-factor solution (Eigenvalues >1), explaining 40% of the variance, and consistent with each item loading strongly on its respective factor. A confirmatory factor analysis estimating two intercorrelated factors ($r = .593$) showed better model fit than a model estimating one single factor, as indicated by a drop in root mean square error of approximation (RMSEA) and Akaike's Information Criterion (AIC) coefficients (RMSEA .03 vs. .05, AIC 93 vs. 105), confirming that the two scales capture two related but distinct constructs.

## Procedure

The experiment was introduced as a study about early detection of diseases through screening. Participants were provided with a short explanation of what screening is and were given several examples of prominent screening procedures (e.g., for cancer, high cholesterol, etc.). Participants then answered several questions about their beliefs about screening in general and filled in the science literacy test and the numeracy test. Then they received information about the disease, the screening procedure for this disease, and the benefits and harms from screening. Finally, participants answered the questions assessing emotional reactions, comprehension, perceived benefits and harms from screening, and intentions to get screened.

## Analysis

The goal of our analysis was to investigate whether and how the description of the disease, individual differences in science literacy and numeracy, a priori screening beliefs, and the effectiveness of screening influenced perceptions of screening and screening intentions. In particular, we investigated whether these variables influenced perceptions and intentions via promoting or impeding comprehension of evidence.

In the analyses, we first examined simple effects of the experimental manipulations (disease description and screening statistics) on the relevant dependent measures. In particular, we checked whether the description of the disease had the intended effect on the perceived severity of the disease, as well as on emotions, comprehension, perceptions of screening, and screening intentions. We also investigated whether the screening statistics manipulation had a direct effect on comprehension, perceptions of screening, and screening intentions. We then examined correlations between the three types of factors—cognitive skills, emotions, and beliefs—and deliberation time, comprehension, perceptions of screening, and screening intentions.

Finally, we conducted a path analysis in SPSS AMOS to test both direct and indirect effects in a multiple regression framework. We used General Least Squares estimation based on 500 bootstrap samples and calculated bias-corrected 95% confidence intervals (CI). To assess model fit we consulted the root mean square of approximation (RMSEA), the $\chi^2$ test, and the Bayesian Information Criterion (BIC). In particular, RMSEA $< .05$ and a nonsignificant $\chi^2$ test ($p > .05$) were used to assess the overall goodness of fit. The BIC was used for model selection because it takes into account both the statistical goodness of fit and the number of estimated parameters and imposes a strict penalty for increasing the number of parameters (Burnham & Anderson, 2002). In particular, a drop in BIC indicates an improvement in the model. We adopted a standard $\alpha$ level of .05 for all statistical decisions.

# Results

## Experimental Manipulations

**Disease description.** Participants who read the severe disease description rated Greene's disease as more severe ($M = 5.41$, $SD = 1.13$) compared with participants who received the neutral description ($M = 5.15$, $SD = 1.07$), $t(345) = -2.24$, $p = .026$, Cohen's $d = .24$, showing that the manipulation was successful. Participants who read the severe description also reported stronger emotional reactions to the disease ($M = 4.96$, $SD = .99$) compared with participants who read the neutral description ($M = 4.72$, $SD = .84$), $t(345) = -2.42$, $p = .016$, Cohen's $d = .26$. The description had no direct effect on any of the other dependent measures, $t$s $< 1$, $p$s $> .5$.

**Screening statistics.** Participants understood a similar proportion of the information regardless of whether screening was effective ($M = 4.14$, $SD = 2.09$) or not ($M = 4.24$, $SD = 2.14$), $t(345) = .48$, $p = .633$, Cohen's $d = .05$. Those who received information that screening saved lives perceived more benefit from screening ($M = 4.33$, $SD = 1.12$) compared with those who received information that screening saved no lives ($M = 3.78$, $SD = 1.41$), $t(345) = -4.04$, $p < .0001$, Cohen's $d = .44$. Those who received information that screening saved lives also had stronger intentions to get screened ($M = 3.66$, $SD = 1.18$ vs. $M = 3.36$, $SD = 1.38$), $t(345) = -2.11$, $p = .035$, Cohen's $d = .23$. There were no other significant effects, $t$s $< 2$, $p$s $> .2$. Analyses of variance (ANOVA) showed that there were no significant interactions between screening statistics and the disease description on any of the mediator and outcome variables ($p > .05$).

## Correlation Analyses

Correlations between the individual difference measures and the dependent variables are shown in Table 1. Screening effectiveness might moderate some of the relationships shown in Table 1 (e.g., emotions may be more strongly related to intentions to get screened when screening is effective). Hence, before proceeding

Table 1
*Correlations Between Individual Difference Measures (Cognitive Skills, Emotions, and Beliefs) and Dependent Variables*

| Measures | Science literacy | Numeracy | Emotions | Deliberation | Comprehension | Perceived benefits | Perceived harms | Screening intentions |
|---|---|---|---|---|---|---|---|---|
| A priori screening beliefs | .119* | −.023 | .136* | .067 | .004 | .295* | .044 | .228* |
| Science literacy | | .353* | .249* | .475* | .387* | .028 | .022 | −.082 |
| Numeracy | | | −.005 | .261* | .300* | .035 | −.053 | −.063 |
| Emotions | | | | .253* | .208* | .124* | .153* | .095 |
| Deliberation | | | | | .570* | −.081 | .049 | −.151* |
| Comprehension | | | | | | −.166* | .014 | −.239* |
| Perceived benefit | | | | | | | −.123* | .751* |
| Perceived harm | | | | | | | | −.187* |

* $p \leq .05$.

with the path analysis, we checked whether any of the above-mentioned significant relationships between the three factors (cognitive skills, emotions, and beliefs) and deliberation time, comprehension, perceived benefit of screening, and screening intentions depended on whether screening was effective or not. We centered the independent variables and computed the interaction terms between the variable screening effectiveness and the respective factor. Linear regression analyses showed no significant interactions for any of the above-mentioned relationships (all *p*s $>$ .05), suggesting that they did not vary as a function of screening effectiveness.

## Path Analysis: Direct and Indirect Effects

Based on the correlation results (see Table 1), we updated our conceptual model outlined in the introduction by allowing for covariance between (a) science literacy and numeracy and (b) science literacy and a priori screening beliefs (Model 1, a graphical illustration is found in the Appendix).

Table 2 shows the fit indices for Model 1 and all subsequent model comparisons. Model 1 showed poor overall fit. To improve fit, we started with fixing the nonsignificant paths to 0. In Model 2, all path weights from Model 1 that had an absolute value of the $\beta$ weight $<$.1 and $p > .05$ were dropped (i.e., fixed to 0), which improved model fit considerably. Next, we consulted modification indices. We considered freeing a parameter if the value of the index was $>$4, and the proposed relationship was theoretically meaningful. The largest modification index was for an effect of science literacy on emotional reactions, that was consistent with previous research documenting links between numeracy and emotional reactions (Peters et al., 2006; Petrova et al., 2014). Indeed, freeing this parameter improved model fit (see Model 3, Table 2). Next, we allowed for two changes that were consistent with the theoretical model of the "affect heuristic," stating that affective reactions toward an activity determine the perceived benefits and perceived risks associated with that activity (Slovic, Finucane, Peters, & MacGregor, 2004), and that these perceived benefits and risks are inversely related (Finucane, Alhakami, Slovic, & Johnson, 2000). In particular, we allowed for an effect from emotional reactions to perceived benefit (Model 4). Finally, we allowed for an effect of perceived harms on perceived benefits, which had a larger modification index (MI = 11) than an effect of perceived benefits on perceived harms (MI = 8; Model 5). Both modifications improved model fit and the final Model 5 showed excellent overall goodness of fit. Figure 2 shows results from the final Model 5 that explained 60% of the variance in screening intentions.

**Cognitive skills.** Both science literacy, standardized effect (*SEF*) = .328[.225, .421], $p$ = .004, and numeracy, *SEF* = .208[.111, .316], $p$ = .004, had total positive effects on comprehension of screening statistics. Both science literacy, *SEF* = .12 [.015, .221], $p$ = .035, and numeracy, *SEF* = .15 [.055, .251], $p$ = .004, had significant direct effects on comprehension. In addition, science literacy was indirectly related to comprehension via both emotional engagement and deliberation time (see Figure 2). In contrast, numeracy was indirectly related to comprehension via deliberation time only (see Figure 2).

**Emotions.** There was a marginally significant indirect effect of the disease description on screening intentions, *SEF* = .008[95% CI −.001, .021], $p$ = .071. Actually, the disease description had two distinct, opposite effects on intentions, which could explain the small total effect. The severe disease description was related to stronger emotional reactions (see Figure 2). On one hand, partially consistent with H2a (i.e., that emotions will be associated with worse comprehension and stronger screening intentions), these heightened emotional reactions were related to more perceived benefit of screening, *SEF* = .196, [.053, .372], $p$ = .015, and to stronger screening intentions, *SEF* = .719, [.653, .777], $p$ = .004, regardless of comprehension. On the other hand, consistent with H2b (i.e., that emotions will be related to better comprehension and weaker screening intentions), these heightened emotions were related to more time spent reading the information, which was related to better comprehension, *SEF* = .071, [.024, .133], $p$ = .004, and weaker screening intentions, *SEF* = −.284, [−.381, −.191], $p$ = .004.

**Beliefs.** Consistent with H3b (i.e., that beliefs will be related to stronger intentions regardless of comprehension), a priori screening beliefs were related to increased screening intentions via stronger perceptions of benefit, *SEF* = .212[.124, .293], $p$ = .004, but had no effect on comprehension.

Finally, the statistical information that screening saved lives was related to stronger screening intentions via stronger perceptions of benefit, *SEF* = .151[.085, .224], $p$ = .004 (see Figure 2). Moreover, although perceived harm from screening was related to fewer perceived benefits and intentions to participate in screening, none of the variables in the model were related to perceived harms. Perceived harms were also a worse predictor of screening intentions than perceived benefits ($\beta_{harms} = -.11$ vs. $\beta_{benefits} = .72$, see Figure 2).

Table 2
*Indices for Model Comparison*

| Model | RMSEA | BIC | *df* | $\chi^2$ | *p* | Modification |
|---|---|---|---|---|---|---|
| 1 | .07 | 296 | 27 | 68 | <.001 | None |
| 2 | .05 | 233 | 39 | 76 | <.001 | Drop nonsignificant paths with $p > .05$ and $\beta < .1$. |
| 3 | .04 | 221 | 38 | 58 | .021 | Estimate path from science literacy to emotional reactions. |
| 4 | .03 | 218 | 37 | 49 | .095 | Estimate path from emotional reactions to perceived benefit. |
| 5 | .02 | 214 | 36 | 39 | .329 | Estimate path from perceived harm to perceived benefit. |

*Note.* RMSEA = root mean square error of approximation; BIC = Bayesian Information Criterion. Smaller values of BIC indicate a better model. Model 1 = initial model. Nonsignificant paths dropped in Model 2 were: all effects on perceived harm, effects of a priori screening beliefs on deliberation, comprehension, and screening intentions; effects of screening statistics on comprehension and intentions; effects of science literacy on perceived benefits, and emotional reactions on comprehension.

Table 3
*Comprehension Items: Questions Used to Assess Comprehension of Information About Benefits and Harms From Screening for Greene's Disease*

| Question text | Answer options | Correct | % correct |
|---|---|---|---|
| 1) For some people screening for Greene's disease can be more harmful than beneficial. | (a) True (b) False (c) I don't know | (a) | 72 |
| 2) Screening for Greene's disease has only benefits for everyone who participates. | (a) True (b) False (c) I don't know | (b) | 71 |
| 3) Screening for Greene's disease can detect abnormal cells that do not present any danger to the person. | (a) True (b) False (c) I don't know | (a) | 55 |
| 4) For each 1,000 people who participated in screening for Greene's disease, how many lives were saved by screening? | ____ people of 1.000 | 2/0[a] | 21 |
| 5) Is the number of people saved by screening larger than the number of people who will undergo unnecessary treatment? | (a) Yes (b) No (c) Both numbers are equal (d) I don't know | (b) | 40 |
| 6) If 2,000 people participate in screening, how many of them will die from Greene's disease in the next 5 years? | ___ people of 2.000 | 6/10[a] | 58 |
| 7) If 2,000 people *do not* participate in screening, how many of them will die from Greene's disease in the next 5 years? | ___ people of 2.000 | 10 | 60 |
| 8) Screening reduces the risk of dying from Greene's disease by ___%. | (a) 0 (b) 20 (c) 40 (d) 60 (e) 80 (f) 100 | (c)/(a)[a] | 43 |

*Note.* % correct = percentage of participants who answered correctly. Adapted from Petrova, Garcia-Retamero, and Cokely (2015).
[a] Depending on the experimental condition: Effective/not effective.

## Discussion

Higher numeracy and science literacy independently predicted better comprehension, providing support for H1. A more severe disease description was related to more worry and fear from the disease. These heightened emotions had both beneficial and detrimental effects on comprehension and intentions to get screened. On one hand, stronger emotional reactions were related to more perceived benefits and stronger screening intentions; however, this effect was independent of comprehension, providing only partial support for H2a (i.e., that emotions would be related to worse comprehension and stronger screening intentions). On the other hand, consistent with H2b, stronger emotional reactions were related to longer deliberation time, better comprehension, and weaker screening intentions. Finally, our hypothesis that stronger positive a priory beliefs about screening would be related to worse comprehension (H3a) was not supported. Instead, positive a priori beliefs were associated with stronger intentions to get screened, regardless of the evidence of screening effectiveness and its com-

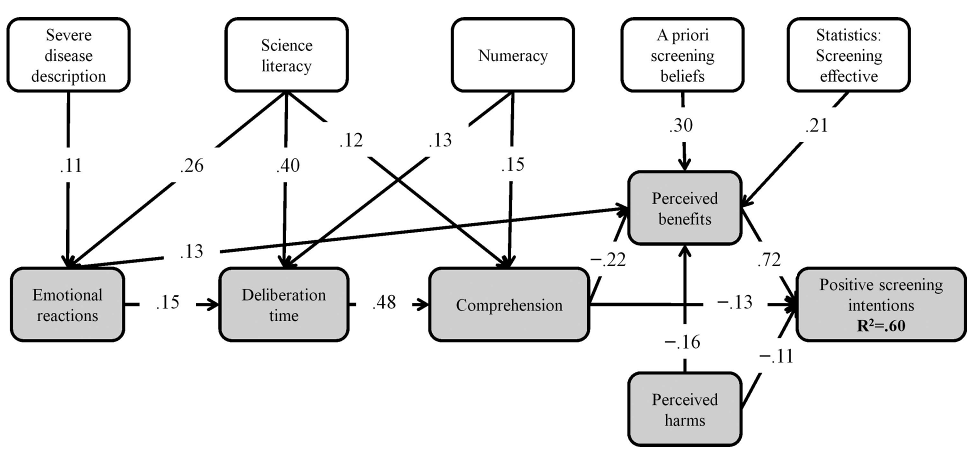


*Figure 2.* Path model of the effects of cognitive skills, emotions, and beliefs on comprehension, perceptions, and screening intentions (Model 5). Coefficients are standardized β weights. Only significant paths are displayed ($p < .05$). There were significant covariances between science literacy and numeracy (.34) and science literacy and a priori beliefs (.14).

prehension, a result that is consistent with H3b. Finally, perceived benefits were more predictive of intentions than perceived harms.

## Cognitive Skills

Several studies have shown the beneficial effects of numeracy on risk comprehension in the context of screening (Davids et al., 2004; Galesic, Garcia-Retamero, & Gigerenzer, 2009; Lipkus et al., 2010; Reyna et al., 2009; Schwartz et al., 1997). However, to our knowledge this is the first study that shows that science literacy has a comparable, independent effect on comprehension. In this study, science literacy was operationalized roughly as knowledge of basic scientific methods used to derive conclusions (e.g., the belief that a control group is needed to establish treatment effectiveness and that correlation does not imply causation; National Science Foundation, 2014). This knowledge is vital for evaluating benefits and harms from screening (e.g., knowing what comparison is important). It is also theoretically different from the ability to calculate risks and proportions captured by numeracy, although both concepts share some variance (Kahan et al., 2012; Schwartz, Woloshin, & Welch, 2005).

In addition, the path analysis showed that the two abilities influenced comprehension in slightly different ways. While both abilities had direct effects on comprehension, their indirect effects differed: Numeracy was indirectly related to comprehension via deliberation time only, and science literacy was related to comprehension via both emotional reactions and deliberation time. In other words, the difference was that participants' numeracy was not related to their emotional reactions to the disease. This result may appear surprising at first sight, given that previous studies have related numeracy to affective reactions (Peters et al., 2006; Petrova et al., 2014). However, these studies have shown that numeracy is related to more precise feelings about numbers and higher emotional sensitivity to risks. That is, individuals with high numeracy have been often found to derive affective meaning from number comparisons (e.g., more differentiated feeling to 5% vs. 10% risk of disease, Peters et al., 2006; Petrova et al., 2014), something that was not made salient in the disease description. Instead, what was made salient were the characteristics and severity of the disease, which affected participants' emotions. It is then possible that higher science literacy contributed to deriving more meaning from the information about the nature of the fictitious disease and its treatment (e.g., because of better understanding of medical terms or knowledge of similar diseases). This can explain why participants with higher science literacy perceived the disease as more frightening on average.

The indirect effect of science literacy on comprehension via deliberation time is in accordance with the hypothesis that participants with higher science literacy may be more likely to spend time processing the information (e.g., because of interest or familiarity with such type of information). In that sense, much like the way numeracy scales capture multiple numerical competencies (Peters & Bjalkebring, 2015), the science literacy scale may capture not only scientific knowledge per se but also people's need for cognition (i.e., their inclination toward effortful cognitive abilities; Cacioppo & Petty, 1982). Future research can investigate whether need for cognition and knowledge of scientific methods affect judgment and decision making independently.

Finally, although the science literacy items we used showed enough discriminability and good results despite a ceiling effect, the field could benefit from a more extensive validated and reliable measure of scientific reasoning (e.g., Drummond & Fischhoff, 2015). For example, the science literacy scale contains some numeracy-like items (see Items 2 and 3 in the Appendix). In addition, the science literacy items were relatively easy and the numeracy items relatively difficult. This means that different performance on the two scales may not necessarily indicate differences in the underlying constructs but in the different difficulty levels. This limitation can be overcome in future research by using scales of comparable difficulty.

Note that the time spent on the page with statistical information is only a rough proxy for deliberation. Deliberation time may partially reflect one's perceived self-efficacy to comprehend the information, as well as one's experience and liking for this kind of information. The deliberation measure could also reflect an early selection metacognitive process, such that individuals who score high on cognitive abilities are more likely to process the information thoroughly and provide better monitoring during subsequent tasks (e.g., early selection vs. late correction; Cokely & Kelley, 2009; Ghazal et al., 2014). This might have helped these people find the correct answers once the questions were revealed. Alternatively, one could expect that people with higher cognitive abilities would process the information more efficiently and would actually spend less time on the page with information. Our results, however, do not support this alternative hypothesis. Nevertheless, although we think that the deliberation measure most likely reflects interest or more thorough strategic processing (see also Cokely & Kelley, 2009; Ghazal et al., 2014), it could also reflect conscientiousness or effort invested in the experiment in general. One way to deal with the limitations of this proxy measure in future research is to use process tracing methods like think-aloud protocols, which can offer a more direct insight into people's deliberation.

## Emotions

Evoking stronger negative emotional reactions toward the disease had two independent, opposite effects on screening intentions. On one hand, it was related to longer deliberation time and better comprehension, thereby decreasing intentions to get screened. One could say that in this case emotions had a *beneficial* effect on decision making, potentially facilitating informed decision making though increasing motivation and comprehension. This effect is also in line with the proposed function of affect as a "spotlight" or "motivator" (Peters et al., 2006). On the other hand, evoking stronger negative emotional reactions toward the disease was related to more perceived benefit of screening, thereby increasing intentions to get screened. More important, this effect persisted regardless of comprehension and the effectiveness of screening, which were controlled for in the model. This means that participants who were more afraid of the disease perceived more benefit, even from screening that did not reduce chances of dying. In this case, this effect of emotions could be considered *detrimental* and leading to inferior, "noninformed" decision making. We should note that stronger emotional reactions were not related to lower comprehension, suggesting that they would not interfere with informed decision making by influencing how information is processed (e.g., paying less attention to numerical information and failing to answer the comprehension questions correctly). Rather, individuals who were more worried about the disease might have given more weight to the benefits of screening or followed a heuristic decision strategy (e.g., "always adopt a preventive behavior," "early detection is always good"). These results are in line with research on financial and medical decisions showing that when the consequences of an event are emotionally powerful, decision makers are less sensitive to the exact probability of its occurrence (in this case the probability of benefit or harm from screening; Lejarraga, Pachur, Frey, & Hertwig, 2015; Pachur et al., 2014; Petrova et al., 2014; Rottenstreich & Hsee, 2001). Instead, they are more concerned about avoiding the focal negative outcome (in this case death from the disease). These results are also in line with a bulk of evidence suggesting that people process risk and benefit information not only cognitively but also emotionally, and that emotions are often even more influential in decisions about health risks (Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2004; Zikmund-Fisher et al., 2010).

## Beliefs

Much like the detrimental effect of emotions on intentions, positive a priori screening beliefs were related to more perceived benefit from screening and stronger intentions to get screened, regardless of comprehension and screening effectiveness. However, positive a priori screening beliefs had no effect on comprehension. Rather individuals who had more positive beliefs about screening (e.g., people who thought that screening was always beneficial or that foregoing screening was irresponsible) might have given more weight to the benefit of screening, even when there was statistically none. For instance, they could have given more weight to other potential benefits not captured by mortality reduction (e.g., the peace of mind in case of a negative screening result).

Our results showed that both strong a priori beliefs and emotional reactions are related to more perceived benefit of screening. This suggests that participants may have engaged in motivated reasoning (Kunda, 1990). Often when people are emotionally invested in a topic, they can trust evidence selectively in patterns that promote their goals and support their expectations; they can also discount or dismiss information that would cause them to experience dissonance or anxiety, especially when it is contrary to their beliefs (Kahan, 2013; Lewandowsky et al., 2012). This suggests that some participants (e.g., those who were very worried about the disease and those who believed screening was desirable) may have been distrustful of the counterintuitive information about harms and lack of benefits. They could also have dismissed it altogether at the moment of decisions making because it did not validate their beliefs that screening is beneficial and did not relieve their fears from a severe disease. This potential mechanism may be even stronger in out-of-the-lab real screening decisions. If this is the case, previous research suggests that it may be difficult to readjust people's beliefs about the inherent goodness of screenings that have been reinforced by numerous screening campaigns (Lewandowsky et al., 2012).

National surveys show that many individuals share the positive a priori screening beliefs assessed in this study: they believe that screenings are almost always beneficial, overestimate their benefits, are unaware of potential harms, and perceive screening as an obligation to family and society (Gigerenzer et al., 2009; Hersch et al., 2013; Hoffman et al., 2010; Schwartz et al., 2004, 2000; Waller et al., 2015). The results of this study suggest that such beliefs can to an extent translate into greater willingness to get screened, even when the evidence is communicated in a simple, user-friendly way, and does not show clear benefits of screening. More broadly, persuasive screening campaigns encouraging screenings without specifying the extent of benefit or mentioning possible harms could contribute to "nonevidence-based" decision making, even when the evidence is provided. Future research should address this possibility with a more ecological approach.

## Perceived Benefits Versus Perceived Harms

Perceived benefits were central in the model, while perceived harms were less predictive of intentions and largely unrelated to cognitive skills, emotions, or beliefs. This suggests that in the context of screening or prevention in general "potential" motivation rather than "security" motivation may be the main driver behind decisions. In her seminal work in the 1980s, Lola Lopes proposed that risk-averse individuals are more motivated by security (e.g., they weigh the worst outcomes more heavily), while risk-seeking individuals are more motivated by potential (e.g., they weigh the best outcomes more heavily; Lopes, 1987). The current results suggest that the importance of these motivations could be also context-dependent, such that when it comes to prevention (i.e., a gain context) people give more weight to potential benefits than to harms (see also Garcia-Retamero & Cokely, 2011). For instance, consistent with this context dependence, in the context of insurance decisions against a potential loss, negative emotions (e.g., fear) were more predictive of decisions than positive emotions (Petrova et al., 2014).

## Limitations

Although we explained the concept of overtreatment to participants and referred to it as a harm resulting from screening, participants may not have considered overtreatment when answering the perceived harms questions. In other words, although participants may have considered the (unnecessary) treatment as very harmful, they may have failed to consider it as a consequence of screening (e.g., because given a positive screening result one can still choose not to undergo treatment). Consistent with this proposition, previous research suggests that some people see overtreatment as an issue of follow-up care rather than screening participation (Waller et al., 2013) and when naming harms from screening people more often focus on harms from the screening procedure itself rather than harms further along the screening cascade (Sutkowi-Hemstreet et al., 2015). Perhaps if the screening procedure had been described as invasive, risky, or costly, perceived harm from screening would have shown stronger association with the other variables in the model. It is also possible that participants were simply very accepting of harms from overtreatment (i.e., did not consider it a deterrent to screening participation). Consistent with this, a probabilistic national survey in the United States found that 56% of respondents would want to be screened for pseudodisease (i.e., cancers that are slow growing or harmless and would not threaten the person's life in their lifetime; Schwartz et al., 2004). Finally, another possibility is that the made-up laser treatment, even in the high severity condition, was not perceived to be as invasive as the traditional cancer treatments like chemotherapy or surgery, which may be more likely to elicit decision-relevant perceptions of harm. Future research should investigate in more detail how lay perceptions of harms from screening are formed and in what way they are important for screening decisions.

In this study we measured comprehension with a set of items that assessed (a) understanding of the counterintuitive idea that screening can be harmful for some individuals, and (b) understanding of the benefits and harms of screening using numerical information. Most of the items in this study measured verbatim comprehension. Future research can extend our research by using items that measure gist comprehension (e.g., the difference between benefits and harms, Elstad et al., 2015). Beliefs and emotions may affect gist comprehension differently or to a greater extent than verbatim comprehension, and gist comprehension may be more predictive of decision than verbatim comprehension (Reyna, 2008, 2012).

Some limitations of this study were that the disease and screening were fictitious, the study population was homogenous, and was comprised of individuals who had little experience with health problems in general or screenings in particular. However, the evidence of benefits and harms from some real-world screenings is still under debate (Barratt, 2015) and is often not communicated to patients in campaigns or in clinical practice (Gigerenzer, 2014; Hoffman et al., 2010; Wegwarth & Gigerenzer, 2011), leaving some real-world decision makers just as naïve to the information as our participants. Moreover, the artificial context allowed us to experimentally manipulate the severity of the disease and the extent of benefit of screening. This would be difficult to achieve for a real disease like cancer, or a real screening program, with which participants have experience and where existing knowledge could influence responses. Also, it is important to keep in mind that the (statistical) information about screening presented to participants in this study may not be representative of all information that experts, patients, and other stakeholders may find relevant for decisions.

Some of the proposed hypotheses were not supported. In particular, we obtained no evidence suggesting that stronger emotional reactions (H2a) or stronger positive a priori screening beliefs (H3a) are related to worse comprehension. Differences between the hypothesized and the final model could reflect that the effects that were not observed had no bearing on screening intentions. However, as with any experimental research, it is possible that the observed model is specific to the scenario and participant population at hand. To build a comprehensive theory of complex screening decisions, our results should be replicated and extended with more diverse populations in more ecological settings.

## Implications for Future Research

The majority of studies investigating screening intentions and adherence have utilized theoretical frameworks that do not give a central role to comprehension of numerical risk and benefit information (e.g., the Health Belief Model, Rosenstock, 1974; and Theory of Reasoned Action, Fishbein, 1980). Instead, to address public health demands and experts' recommendations, most previous research on screening using these models has followed a persuasion-based tradition, looking to eliminate barriers to screening adherence (e.g., Austin, Ahmad, McNally, & Stewart, 2002; Curry & Emmons, 1994; Johnson, Mues, Mayne, & Kiblawi, 2008; Miller, Shoda, & Hurley, 1996). While such persuasion-based models may still apply to beneficial screenings with negligible harms, updated epidemiological evidence shows that many screenings can result in both benefits *and* serious harms (e.g., overdiagnosis in several cancer screenings, Esserman, Thompson, & Reid, 2013; fetal loss in prenatal screening, Lerman, Croyle, Tercyak, & Hamann, 2002; van den Berg, Timmermans, ten Kate, van Vugt, & van der Wal, 2006). For such screenings, policies promoting informed rather than persuasion-based decision making are recommended (Rimer et al., 2004; Sheridan et al., 2004; Woloshin, Schwartz, Black, & Kramer, 2012) and informed choice is a more justified indicator of success than uptake rates (Strech, 2014).

Achieving informed decision making about screening can be challenging because some medical concepts may be difficult to explain and the associated numerical evidence can be confusing. An informative example here is the case of harm from overdiagnosis, which we have simulated in this study. People consider information about overdiagnosis as complicated but important (Hersch et al., 2013; Waller et al., 2013). For example, research in women considering screening with mammography shows that screening intentions may depend heavily on the exact risk of overdiagnosis and that provision of numerical information about both screening benefits and harms from overdiagnosis increases informed choices (Hersch et al., 2013, 2015). While this highlights the importance of providing numerical information in general, it is especially important for future research to address how such information should be tailored so that each patient can benefit from it. The most effective format or the optimal amount of information provided may depend on the patient's specific information needs (Zikmund-Fisher, 2013) and numeracy (Garcia-Retamero & Cokely, 2013; Schwartz, 2011). For example, in a recent experiment a visual aid in the form of an icon array increased comprehension of benefits and harms from two common cancer screenings for many participants. However, it had an opposite effect among some participants who perceived the consequences of a cancer diagnosis as extremely severe (Petrova et al., 2015).

Given that comprehension of benefits and harms is central to informed decision making, we need updated theoretical models that apply to screening procedures where informed decision making is recommended. These should give a more central role to comprehension and predict how it influences decisions (e.g., Fuzzy Trace Theory; Reyna, 2008, 2014) and how it is influenced by relevant person- and situation-based factors. This study identified some factors whose effects may generalize to other decision making contexts where numerical information is involved (e.g., the role of emotions and science literacy). It also identified some issues that may be unique to complex screening decisions. One example is the nonevidence based influence of a priori beliefs about the goodness of screening on intentions. Another one is the unclear, potentially complex antecedents of perceived harm from screening that require further study in a more naturalistic setting.

## References

Austin, L. T., Ahmad, F., McNally, M. J., & Stewart, D. E. (2002). Breast and cervical cancer screening in Hispanic women: A literature review using the health belief model. *Women's Health Issues, 12,* 122–128. http://dx.doi.org/10.1016/S1049-3867(02)00132-9

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior, 31,* 143–164. http://dx.doi.org/10.1177/1090198104263660

Barratt, A. (2015). Overdiagnosis in mammography screening: A 45 year journey from shadowy idea to acknowledged reality. *British Medical Journal (Clinical Research Edition), 350,* h867. http://dx.doi.org/10.1136/bmj.h867

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. New York, NY: Springer.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116–131. http://dx.doi.org/10.1037/0022-3514.42.1.116

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7,* 25–47.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* 20–33.

Curry, S. J., & Emmons, K. M. (1994). Theoretical models for predicting and improving compliance with breast cancer screening. *Annals of Behavioral Medicine, 16,* 302–316.

Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine, 19,* 310–315. http://dx.doi.org/10.1111/j.1525-1497.2004.20801.x

Drummond, C., & Fischhoff, B. (2015). Development and validation of the Scientific Reasoning Scale. *Journal of Behavioral Decision Making*. Advance online publication. http://dx.doi.org/10.1002/bdm.1906

Elstad, E. A., Sutkowi-Hemstreet, A., Sheridan, S. L., Vu, M., Harris, R., Reyna, V. F., . . . Brewer, N. T. (2015). Clinicians' perceptions of the benefits and harms of prostate and colorectal cancer screening. *Medical Decision Making, 35,* 467–476. http://dx.doi.org/10.1177/0272989X15569780

Esserman, L. J., Thompson, I. M., Jr., & Reid, B. (2013). Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *Journal of the American Medical Association, 310,* 797–798. http://dx.doi.org/10.1001/jama.2013.108415

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13,* 1–17. http://dx.doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fishbein, M. (1980). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation, 27,* 65–116.

Fisher, J. D., & Fisher, W. A. (2002). The information-motivation-behavioral skills model. In R. J. DiClemente, R. A. Crosby, & M. C. Kegler (Eds.), *Emerging theories in health promotion practice and research: Strategies for improving public health* (pp. 40–70). San Francisco, CA: Jossey-Bass.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine, 170,* 462–468. http://dx.doi.org/10.1001/archinternmed.2009.481

Galesic, M., & Garcia-Retamero, R. (2011). Do low-numeracy people avoid shared decision making? *Health Psychology, 30,* 336–341. http://dx.doi.org/10.1037/a0022723

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28,* 210–216. http://dx.doi.org/10.1037/a0014474

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied, 17,* 270–287. http://dx.doi.org/10.1037/a0023677

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22,* 392–399. http://dx.doi.org/10.1177/0963721413491570

Garcia-Retamero, R., & Galesic, M. (2010). How to reduce the effect of framing on messages about health. *Journal of General Internal Medicine, 25,* 1323–1329. http://dx.doi.org/10.1007/s11606-010-1484-9

Garcia-Retamero, R., & Galesic, M. (2012). Doc, what would you do if you were me? On self-other discrepancies in medical decision making. *Journal of Experimental Psychology: Applied, 18,* 38–51. http://dx.doi.org/10.1037/a0026018

Garcia-Retamero, R., Müller, S. M., Catena, A., & Maldonado, A. (2009). The power of causal beliefs and conflicting evidence on causal judgments and decision making. *Learning and Motivation, 40,* 284–297. http://dx.doi.org/10.1016/j.lmot.2009.04.001

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology, 33,* 920–928. http://dx.doi.org/10.1037/hea0000061

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9,* 15–34.

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. *British Medical Journal, 348,* g2636. http://dx.doi.org/10.1136/bmj.g2636

Gigerenzer, G., Mata, J., & Frank, R. (2009). Public knowledge of benefits of breast and prostate cancer screening in Europe. *Journal of the National Cancer Institute, 101,* 1216–1220. http://dx.doi.org/10.1093/jnci/djp237

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews, 6,* CD001877.

Grimes, D. A., & Schulz, K. F. (2002). Uses and abuses of screening tests. *Lancet, 359,* 881–884. http://dx.doi.org/10.1016/S0140-6736(02)07948-5

Harris, R. P., Sheridan, S. L., Lewis, C. L., Barclay, C., Vu, M. B., Kistler, C. E., . . . Brewer, N. T. (2014). The harms of screening: A proposed taxonomy and application to lung cancer screening. *Journal of the American Medical Association Internal Medicine, 174,* 281–285. http://dx.doi.org/10.1001/jamainternmed.2013.12745

Hersch, J., Barratt, A., Jansen, J., Irwig, L., McGeechan, K., Jacklyn, G., . . . McCaffery, K. (2015). Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: A randomised controlled trial. *The Lancet, 385,* 1642–1652. http://dx.doi.org/10.1016/S0140-6736(15)60123-4

Hersch, J., Jansen, J., Barratt, A., Irwig, L., Houssami, N., Howard, K., . . . McCaffery, K. (2013). Women's views on overdiagnosis in breast cancer screening: A qualitative study. *British Medical Journal (Clinical Research Edition.), 346,* f158. http://dx.doi.org/10.1136/bmj.f158

Hoffman, R. M., Lewis, C. L., Pignone, M. P., Couper, M. P., Barry, M. J., Elmore, J. G., . . . Zikmund-Fisher, B. J. (2010). Decision-making processes for breast, colorectal, and prostate cancer screening: The DECISIONS survey. *Medical Decision Making, 30*(Suppl.), 53S–64S.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm, P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews, 1,* CD004720.

Johnson, C. E., Mues, K. E., Mayne, S. L., & Kiblawi, A. N. (2008). Cervical cancer screening among immigrants and ethnic minorities: A systematic review using the Health Belief Model. *Journal of Lower Genital Tract Disease, 12,* 232–241. http://dx.doi.org/10.1097/LGT.0b013e31815d8d88

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision Making, 8,* 407–424.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change, 2,* 732–735. http://dx.doi.org/10.1038/nclimate1547

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480–498. http://dx.doi.org/10.1037/0033-2909.108.3.480

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education, 84,* 71–94. http://dx.doi.org/10.1002/(SICI)1098-237X(200001)84:1<71::AID-SCE6>3.0.CO;2-C

Lejarraga, T., Pachur, T., Frey, R., & Hertwig, R. (2015). Decisions from experience: From monetary to medical gambles. *Journal of Behavioral Decision Making*. Advance online publication. http://dx.doi.org/10.1002/bdm.1877

Lerman, C., Croyle, R. T., Tercyak, K. P., & Hamann, H. (2002). Genetic testing: Psychological aspects and implications. *Journal of Consulting and Clinical Psychology, 70,* 784–797. http://dx.doi.org/10.1037/0022-006X.70.3.784

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest, 13,* 106–131. http://dx.doi.org/10.1177/1529100612451018

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,* 464–473. http://dx.doi.org/10.1177/0272989X09360371

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21,* 37–44. http://dx.doi.org/10.1177/0272989X0102100105

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127,* 267–286.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology, 20,* 255–295. http://dx.doi.org/10.1016/S0065-2601(08)60416-5

Miller, S. M., Shoda, Y., & Hurley, K. (1996). Applying cognitive-social theory to health-protective behavior: Breast self-examination in cancer screening. *Psychological Bulletin, 119,* 70–94. http://dx.doi.org/10.1037/0033-2909.119.1.70

Moyer, V. A., & the U.S. Preventive Services Task Force. (2012). Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine, 157,* 120–134. http://dx.doi.org/10.7326/0003-4819-157-2-201207170-00459

National Science Foundation. (2014). *Science and engineering indicators*. Washington, DC: Author.

Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision, 1,* 64–78. http://dx.doi.org/10.1037/dec0000006

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

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. http://dx.doi.org/10.1111/j.1460-2466.2006.00287.x

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17,* 407–413. http://dx.doi.org/10.1111/j.1467-9280.2006.01720.x

Petrova, D., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making, 35,* 847–858. http://dx.doi.org/10.1177/0272989X15587676

Petrova, D., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27,* 191–199. http://dx.doi.org/10.1002/bdm.1803

Reyna, V. F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making, 28,* 850–865. http://dx.doi.org/10.1177/0272989X08327066

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making, 7,* 332–359.

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,* 943–973. http://dx.doi.org/10.1037/a0017327

Rimer, B. K., Briss, P. A., Zeller, P. K., Chan, E. C., & Woolf, S. H. (2004). Informed decision making: What is its role in cancer screening? *Cancer, 101*(Suppl.), 1214–1228. http://dx.doi.org/10.1002/cncr.20512

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior, 2,* 328–335. http://dx.doi.org/10.1177/109019817400200403

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science, 12,* 185–190. http://dx.doi.org/10.1111/1467-9280.00334

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127,* 966–972. http://dx.doi.org/10.7326/0003-4819-127-11-199712010-00003

Schwartz, L. M., Woloshin, S., Fowler, F. J., Jr., & Welch, H. G. (2004). Enthusiasm for cancer screening in the United States. *Journal of the American Medical Association, 291,* 71–78. http://dx.doi.org/10.1001/jama.291.1.71

Schwartz, L. M., Woloshin, S., Sox, H. C., Fischhoff, B., & Welch, H. G. (2000). US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: Cross sectional survey. *British Medical Journal (Clinical Research Edition), 320,* 1635–1640. http://dx.doi.org/10.1136/bmj.320.7250.1635

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2005). Can patients interpret health information? An assessment of the medical data interpretation test. *Medical Decision Making, 25,* 290–300. http://dx.doi.org/10.1177/0272989X05276860

Schwartz, P. H. (2011). Questioning the quantitative imperative: Decision aids, prevention, and the ethics of disclosure. *The Hastings Center Report, 41,* 30–39. http://dx.doi.org/10.1353/hcr.2011.0029

Sheridan, S. L., Harris, R. P., & Woolf, S. H., & the Shared Decision-Making Workgroup of the U.S. Preventive Services Task Force. (2004). Shared decision making about screening and chemoprevention. a suggested approach from the U.S. Preventive Services Task Force. *American Journal of Preventive Medicine, 26,* 56–66. http://dx.doi.org/10.1016/j.amepre.2003.09.011

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24,* 311–322. http://dx.doi.org/10.1111/j.0272-4332.2004.00433.x

Slovic, P., Peters, E., Finucane, M. L., & Macgregor, D. G. (2005). Affect, risk, and decision making. *Health Psychology, 24*(Suppl.), S35–S40. http://dx.doi.org/10.1037/0278-6133.24.4.S35

Strech, D. (2014). Participation rate or informed choice? Rethinking the European key performance indicators for mammography screening. *Health Policy, 115,* 100–103. http://dx.doi.org/10.1016/j.healthpol.2013.11.012

Sutkowi-Hemstreet, A., Vu, M., Harris, R., Brewer, N. T., Dolor, R. J., & Sheridan, S. L. (2015). Adult patients' perspectives on the benefits and harms of overused screening tests: A qualitative study. *Journal of*

*General Internal Medicine, 30,* 1618–1626. http://dx.doi.org/10.1007/s11606-015-3283-9

van den Berg, M., Timmermans, D. R., ten Kate, L. P., van Vugt, J. M., & van der Wal, G. (2006). Informed decision making in the context of prenatal screening. *Patient Education and Counseling, 63,* 110–117. http://dx.doi.org/10.1016/j.pec.2005.09.007

Waller, J., Douglas, E., Whitaker, K. L., & Wardle, J. (2013). Women's responses to information about overdiagnosis in the UK breast cancer screening programme: A qualitative study. *British Medical Journal Open, 3,* e002703. http://dx.doi.org/10.1136/bmjopen-2013-002703

Waller, J., Osborne, K., & Wardle, J. (2015). Enthusiasm for cancer screening in Great Britain: A general population survey. *British Journal of Cancer, 112,* 562–566. http://dx.doi.org/10.1038/bjc.2014.643

Wegwarth, O., & Gigerenzer, G. (2011). "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Education and Counseling, 84,* 251–256. http://dx.doi.org/10.1016/j.pec.2010.07.025

Woloshin, S., Schwartz, L. M., Black, W. C., & Kramer, B. S. (2012). Cancer screening campaigns—Getting past uninformative persuasion. *The New England Journal of Medicine, 367,* 1677–1679. http://dx.doi.org/10.1056/NEJMp1209407

Zikmund-Fisher, B. J. (2013). The right tool is what they need, not what we have: A taxonomy of appropriate levels of precision in patient risk communication. *Medical Care Research and Review, 70*(Suppl.), 37S–49S. http://dx.doi.org/10.1177/1077558712458541

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2010). Risky feelings: Why a 6% risk of cancer does not always feel like 6%. *Patient Education and Counseling, 81*(Suppl.), S87–S93. http://dx.doi.org/10.1016/j.pec.2010.07.041

Received June 26, 2015
Revision received February 8, 2016
Accepted February 14, 2016 ■