Being Below or Above Average: How Presentation Format Affects the Interpretation of Comparative Breast Cancer Risks

Merel D. A. Kersjes

SNR 2014579

Master's Thesis

Communication and Information Sciences

Specialization Communication and Cognition

School of Humanities and Digital Sciences

Tilburg University, Tilburg

Supervisors: R.D. Vromans & dr. N. Bol

Second Reader: H. Engelbrecht

June 2021

Abstract

When communicating comparative breast cancer risks to women, for instance about being below or above average, it is crucial that women have accurate perceptions of their risk and are not disproportionally eased or worried**.** This study examined how comparative breast cancer risks – either below or above average – can best be presented in the context of personalized breast cancer screening; numerical-only or both numerical and visualized. A 2 (risk level: below (7%) vs. above (28%) average) x 3 (presentation format: numerical-only vs. numerical plus bar graph vs. numerical plus icon array) between-subjects design was used. 403 Dutch women between 50 and 75 years old were randomly assigned to one of six hypothetical scenarios about personalized breast cancer screening that contained fictive risk estimates. Accuracy of perceived risk, perceived risk and breast cancer worry were measured both before and after exposure to the scenario. The results showed that presentation format did not influence these three outcome measures. In contrast, risk level affected all three: Accuracy of perceived risk, perceived risk and breast cancer worry were higher in the 28% conditions than in the 7% conditions. Exploratory analyses revealed that the height of one's numeracy affected (accuracy of) risk perceptions and breast cancer worry. These findings implicate that risk level affects both cognitive and affective responses to comparative risk information, but presentation format does not. Future research could test if these effects still hold when presenting actual instead of fictive risk information and in different formats, and how comparative risk information should be presented to less numerate women.

Keywords: risk communication, comparative risk information, visualizations, risk level, perceived risk, accuracy of perceived risk, breast cancer worry

Table of Contents

Being Below or Above Average: How Presentation Format Affects the Interpretation of Comparative Breast Cancer Risks

Imagine you are a 50-year-old woman who received an invitation letter for personalized breast cancer screening. The letter states that screening practices – such as the frequency and modality of screening – will be tailored towards your personal breast cancer risk, which can be determined in a risk assessment. Multiple women in your family, including your mother, have been diagnosed with breast cancer in the past, which makes you assume that you have a higher than average risk of developing breast cancer – a frightening thought. You want to do everything to minimize your chances of getting breast cancer and decide to take part in the risk assessment. You are very curious whether your risk is as high as you assume it is, and if it is higher than other women's. And, when it turns out that your risk is higher than average, what will happen to your personal screening process?

At this moment, breast cancer screening – a strategy aimed at detecting breast cancer at an early stage to reduce mortality and (extensive) treatment – is offered to all Dutch women between 50 and 75 years old (Rijksinstituut voor Volksgezondheid en Milieu, 2020). It is, however, questionable whether the current screening approach, of which the guidelines are designed for women at average risk, maximizes benefits while minimizing harms (e.g., overdiagnosis or overtreatment). An alternative approach that has the potential to optimize the benefit-to-harm ratio is personalized, or tailored, breast cancer screening. Kreuter et al. (1997, 1999) define tailoring as targeting specific people based on individual characteristics that are derived from individual assessment and are related to the relevant outcome. In the case of breast cancer screening, screening practices (e.g., screening interval or technique) would be tailored towards women's personal breast cancer risks – assessed through tools such as the Gail model (Wood et al., 2019) – and preferences, albeit evidence-based (Onega et al., 2014; Román et al., 2019).

When implementing such a personalized screening program, it is crucial that women are well-informed about their personal breast cancer risks. That is, women need to have accurate perceptions of their risk and should not be disproportionately worried or eased. These outcomes, in this study defined as (accuracy of) perceived risk and breast cancer worry, largely depend on how risk information is provided. The context in which risks are presented is crucial. According to Zikmund-Fisher, “Contextual information determines whether data are easy or difficult for people to make sense of and use in decision making” (2019, p. 29). He argues that relevant reference standards are needed to make information meaningful to nonexpert people. In case of risk estimates, Zikmund-Fisher (2019) poses that is impossible to know whether single risks (e.g., 28%) are good or bad without contextual information. People therefore ignore single risks in decision-making (Zikmund-Fisher, 2013; Zikmund-Fisher et al., 2004). Regarding personalized breast cancer screening, this would suggest that breast cancer risks should be presented with relevant contextual information, in this case the average risk of developing breast cancer (i.e., comparative risk).

How comparative risk information is perceived and interpreted, however, might vary between different presentation formats. For example, it has been found that people perceive identical risk estimates to be lower when presented in a bar graph than when presented in an icon array (Schapira et al., 2006). Another study showed that risk estimates in a bar graph are perceived more accurately than only numerically presented risks (Waters et al., 2007). Especially comparative risk information has been found to be influential, in the sense that telling women they are below or above average affects their risk perceptions and subsequent actions (Fagerlin et al., 2007a). Providing comparative risk information should not lead to incorrect over- or underestimations of one’s risk or to excessive worry, since this could negatively affect one’s choices. Although several studies have investigated the effects of different visual formats on cognitive and affective measures (see, e.g., Ghosh et al., 2008, Schapira et al., 2006; Timmermans et al., 2008), no previous study – to my knowledge – included comparative risk information nor differed risk levels (e.g., below or above average). Therefore, it is yet unknown a) which presentation format yields the most accurate interpretations of comparative risks, and b) what role risk level plays.

The aim of the present study is to examine how comparative breast cancer risks should be presented in the leaflets or letters women would receive for personalized breast cancer screening. This knowledge is of crucial importance, since women need to be able to make informed decisions based on accurate risk perceptions. Furthermore, this study aims to fill the gap in research on presenting comparative risks. It will be investigated what effects presentation format (numerical-only or numerical and visualized) and risk level (either below or above average) have on the accuracy and height of women's risk perceptions, as well as on their breast cancer worry. The research question that is dealt with is: How does the presentation format of comparative breast cancer risks affect women's (accuracy of) perceived risk and breast cancer worry, and what role does risk level play?

Theoretical Framework

Lipkus defines risk as "a combined function, often multiplicative, of the probability of loss and consequence of loss (e.g., severity of loss in the physical, psychological, social, and economic realms) (Berry, 2004; Brun, 1994; Slovic, 1999; Thompson & Dean, 1996; Yates & Stone, 1992)" (2007, p. 697). In the context of personalized breast cancer screening, it is about the risk of getting breast cancer and handling according to that risk. When communicating such personal risks to eligible women, it is crucial that they truly understand what their risk means and can make well-informed decisions accordingly. Women should have accurate perceptions of the height of their risk (i.e., not over- or underestimate it), and should not be disproportionally eased or worried about (their risk of) getting breast cancer.

Comparative Risk Information

How people estimate their risk depends largely on whether they solely rate their own risk (i.e., personal risk) or rate it relative to a standard (i.e., comparative risk). That is, showing people that their risk is below or above average affects their cognitive, affective and behavioral responses (Fagerlin et al., 2007a; Klein, 1997). It is argued that the mind uses a dual representation of risk: One system observes the numerical presentation of risk, while the other forms more intuitive responses of the odds of that risk (Windschitl et al., 2002). The former corresponds to personal risk information, while the second system also involves comparative risk information (Schwartz, 2009). For instance, only telling women that they have a 28% chance of getting breast cancer will probably not yield the same cognitive and affective reactions as telling them they have a 28% chance of getting breast cancer while the average risk is 14%, since the latter provides them more information about being at higher than average risk.

Providing comparative risk information seems to resemble the way people think about risks. As explained by Festinger in his social comparison theory (SCT, Festinger, 1954) people have a tendency to compare themselves to others. They find it important how they compare with others (Suls & Wheeler, 2000), especially when these others are similar to themselves (Festinger, 1954). This also seems to be the case in the context of health communication: Previous research has shown that when provided health-related risks, people think about risk in comparative terms, even in the presence of a personal risk estimate (Klein, 1997; Fagerlin et al., 2007a). Therefore, providing comparative breast cancer risk information seems to closely correspond to the tendency to compare oneself to similar others. Comparative risk information directly shows how one relates to others, since it provides information about being at lower or higher than average risk; something that it is unknown when only given one's personal risk.

A study by Fagerlin et al. (2007a), in which women were shown comparative risk information about a hypothetical preventive breast cancer medicine, demonstrates that people think about risks in comparative terms. Although all participants were given the exact same personal risk estimate (6%), the comparative risk estimate varied (either 3% or 12%), implicating that they were either above or below average, respectively. Fagerlin et al. (2007a) found that the former group was more willing to take the medicine than the latter group and that this group was also more likely to believe its risk would be significantly reduced. This implicates that showing people that they are below or above average affects their reactions, even when the personal risk level is the same. While Fagerlin et al. (2007a) kept the personal risk estimate constant and varied the average risk, the present study will keep the average risk estimate constant while varying personal risk level. This corresponds better with the real-world situation, since it is more common that the average risk estimate is constant while one's personal risk level varies.

Varying personal risk levels, such as below or above average risks, are argued to evoke different cognitive and affective responses. French et al. (2004) showed that people who received favourable comparative risk information (i.e., lower personal risk than average risk) perceived their risk as lower and of less magnitude than people who received unfavourable comparative risk information (i.e., higher personal risk than average risk). Higher risk perceptions may lead to more breast cancer worry, since perceived risk and breast cancer worry are positively correlated (Lipkus et al., 2005). McCaul et al. (2003) proved that comparative risk information impacts breast cancer worry, such that women who overestimate their risk have more breast cancer worry. A reason for the varying effects of being below or above average could be that comparative risks are less abstract than personal risks, since women can directly compare themselves to others, which they cannot do with personal risks (Lipkus et al., 2005). It is important to note that providing comparative risk information about being below or above average could improve or impair risk perceptions, possibly depending on women's baseline risk perceptions (Schwartz, 2009). That is why perceived risk and breast cancer worry are measured both before (i.e., baseline measurement) and after exposure to comparative risk information.

Based on the findings of Fagerlin et al. (2007a) and French et al. (2004), and on the impact comparative risk information has on cognitive and affective responses to risks, it is hypothesized that the 28% risk level condition – which is higher than the average risk (i.e., 14%) – leads to higher perceived risk and more breast cancer worry than the 7% risk level condition, which is lower than average. Therefore, the first hypothesis is:

H1: Presenting an above average risk estimate (i.e., 28%) leads to higher perceived risk and more breast cancer worry than presenting a below average risk estimate (i.e., 7%).

Presenting Comparative Risks

How risks are presented affects both cognitive and affective perceptions, since many individuals do not have stable opinions about their personal risks beforehand (Lichtenstein & Slovic, 2006). Presenting risk estimates in a visual display could enhance understanding (see, e.g., Garcia-Retamero & Cokely, 2017; Lipkus, 2007; Spiegelhalter, 2017; Zipkin et al., 2014). Visual displays can summarize substantial amounts of data, can directly show mathematical operations (e.g., subtraction between two bars in a bar graph), can keep the attention of the viewer, and are useful when presenting part-to-whole relationships (e.g., icon arrays portraying x out of 100 people) (Cleveland & McGill, 1984; Lipkus & Hollands, 1999; Reyna & Brainerd, 2008).

Multiple studies found that the addition of visualizations to numerical presentation formats leads to more accurate understanding of the risk information presented (see, e.g., Garcia-Retamero & Galesic, 2010; Garcia-Retamero & Hoffrage, 2013). This could be explained through the dual-coding theory (DCT, Clark & Paivio, 1991). The DCT poses that verbal and visual information are processed separately in working memory. The verbal system processes text, whereas the visual system processes graphics and visuals. When exposed to both verbal and visual information, viewers have to form an internal verbal representation, an internal visual representation, and an internal connected representation of visual and verbal information (Mayer & Sims, 1994). Processing information both verbally and visually leads to better understanding and better integration with prior knowledge (Clark & Paivio, 1991). Regarding comparative risks, additional visualizations directly show if one's below or above average (e.g., one bar higher than the other), whereas providing only numerical information still requires cognitive effort (e.g., interpreting 7% as being lower than 14%). Therefore, providing additional visualizations would enhance processing and potentially improve subsequent interpretations of risk information.

Different Types of Visualizations

Although research suggests that the addition of visualizations improves the interpretation of communicated risks, there seem to be differences between the graphical formats in which risks are presented. Two of the most commonly used visual formats are bar graphs and icon arrays. In icon arrays, icons can be presented in multiple ways, each leading to different mental processing (Zikmund-Fisher et al., 2014b). Especially restroom icons seem to be effective: They are easier to recall, more preferred, and more likely to result in accurate risk perceptions than other icons (Zikmund-Fisher et al., 2014b). A focus group study of Schapira et al. (2001) corroborated that women considered human-like icons to be more meaningful, easier to understand, and easier to identify than bar graphs. Especially regarding comparative risk information, this affective response could be highly influential. When women perceive the comparative risk information to be more meaningful and easier to understand and recall, they might be less likely to incorrectly over- or underestimate their personal risk compared to the average risk.

Previous studies indeed showed that icon arrays and bar graphs, or other formats, have differential effects on cognitive and affective measures. Schapira et al. (2006) compared bar graphs with icon arrays and found that numerically identical risk estimates were perceived as lower when presented in a bar graph compared with an icon array. A similar result was shown in a focus group study on how risks should be communicated, but only for less educated women (Schapira et al., 2001). Moreover, Timmermans et al. (2008) showed that icon arrays had more affective impact than numerically presented risks (i.e., either frequencies or percentages), and risks were perceived as higher when presented in icon array format than when presented numerically.

Based on the outcomes of Schapira et al. (2001, 2006) and Timmermans et al. (2008), it is expected that there is an interaction between risk level and presentation format. As previously stated in the first hypothesis, the above average risk level (28%) is expected to yield the highest risk perceptions and most breast cancer worry. It is hypothesized that icon arrays will strengthen this effect, leading to highest risk perceptions and most breast cancer worry. Therefore, the second hypothesis is:

H2: The effect of presenting an above average (i.e., 28%) risk estimate is strengthened by presentation format, such that numerical plus icon array format leads to higher perceived risk and more breast cancer worry than numerical plus bar graph or numerical-only format.

Besides risk perceptions and breast cancer worry, presentation format also has been found to affect the accuracy of risk perceptions. Brown et al. (2011) showed that, although participants initially preferred bar graphs, they interpreted this type of graph least accurate, whereas interpretation was most accurate in the icon array condition. The positive effects of icon arrays on the accuracy of risk perceptions might have to do with the fact that icon arrays portray frequencies. Frequency formats (e.g., 14 out of 100) seem to be more accurately perceived than probability formats (e.g., 14%) (see, e.g., Garcia-Retamero & Hoffrage, 2013; Hoffrage et al., 2000; Hoffrage & Gigerenzer, 1998). Although multiple studies found that icon arrays lead to highest accuracy of risk perceptions (Brown et al., 2011; Ghosh et al., 2008), none of these studies included comparative risk information nor varied risk level. Therefore, it is yet unknown if presentation format yields the same or different results regarding comparative risk information.

Based on the theory on icons and frequencies, and on the findings of previous studies (Brown et al., 2011; Ghosh et al., 2008; Hawley et al., 2008), it is expected that visually presented risk information leads to more accurate risk perceptions than numerical-only presented risks, with icon arrays leading to most accurate perceptions. Although none of the aforementioned studies included comparative risk, it is expected that, based on their findings:

> **H3:** Accuracy of perceived risk is higher when presenting comparative risk information in a numerical plus visual presentation format compared to numerical-only format, with numerical plus icon array format leading to highest accuracy of perceived risk.

Since risk level has not been investigated yet in the context of visually presented comparative risk information, no *a priori* hypotheses were formulated for risk level. It will be explored if and how risk level affects the relationship between presentation format and accuracy of perceived risk.

Numeracy

The way numerical and graphical risk information are interpreted not only depends on the presentation format, but also on one’s numerical skills, referred to as numeracy. Numeracy is defined as “the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatical, and probabilistic health information needed to make effective health decisions" (Golbeck et al., 2005, p. 375). Brown et al. (2011) showed that numeracy is strongly linked to graphicacy; one's ability to understand and present information in graphical formats (Aldrich & Sheppard, 2000). Numeracy hence affects the interpretation of numerical risks as well as the interpretation of graphs. Also, Ghosh et al. (2008) did not include numeracy in their study on the effects of graphical format on accuracy of perceived risk, but they acknowledge that numeracy could play a role and thus needs to be controlled for in future research. For these reasons it is meaningful to include numeracy as a covariate in the present study.

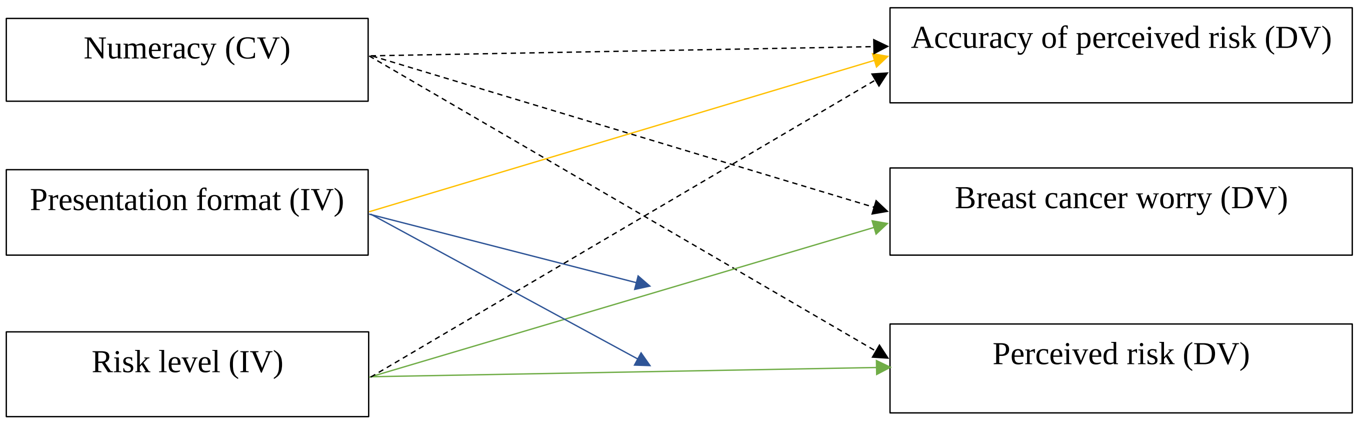
Individual differences in numerical skills have been shown to explain differences in the interpretation of risk information. For example, less numerate individuals seem to be less accurate when estimating the reduction in breast cancer death from screening (Schwartz et al., 1997) and have less accurate perceptions of their health risks (Zikmund-Fisher et al., 2007). Peters et al. (2006) found that highly numerate individuals are less susceptible to framing effects than less numerates, and that they have stronger and more precise affective responses to numerical risk information. Because of these differences it is argued that less numerate individuals benefit mostly from the addition of graphical displays (Lipkus, 2007; Schwartz et al., 1997). More specifically, less numerates seem to benefit mostly from icon arrays (Galesic et al., 2009). Since many studies show that individual differences in numeracy explain variance in the interpretation of risk information and the accuracy of (risk) perceptions, an exploratory subgroup analysis with less and highly numerate women will be conducted, next to the inclusion of numeracy as a control variable.

Conceptual Model

The conceptual model of the present study is depicted in Figure 1. The hypotheses are indicated with colours. The variables in this conceptual model, together with the aforementioned research question and hypotheses[1], were pre-registered within the Open Science Framework (https://osf.io/gk4rz).

Figure 1

The Conceptual Model of the Present Study



Note. CV = control variable, IV = independent variable, DV = dependent variable. Green arrows = H1, blue arrows = H2, yellow arrow = H3, dotted arrow = exploratory analysis.

Method

Design

In this experimental study, a 2 (risk level: below average (7%) vs. above average (28%)) x 3 (presentation format: numerical-only vs. numerical plus bar graph vs. numerical plus icon array) between-subjects design was used. The independent variables were manipulated via hypothetical scenarios. Due to ethical considerations, participants were not given their actual personal breast cancer risk, but were rather given hypothetical scenarios containing a fictive risk level. Hypothetical scenarios are a commonly used method in risk communication research (see, e.g., Fagerlin et al., 2007a; Hawley et al., 2008; Schapira et al., 2006).

---

[1] The pre-registration (https://osf.io/gk4rz) included only two hypotheses, while this study's final version includes three hypotheses. One of the hypotheses was split in two, and the hypotheses are formulated slightly differently from the pre-registration. The content of the hypotheses did not change, nor did the conceptual model; only sequential and textual changes were made.

Dependent variables were ‘perceived risk’, ‘accuracy of perceived risk’ and ‘breast cancer worry’. Furthermore, ‘numeracy’ was considered as a control variable for all dependent variables.

Participants

Dutch women between 50 and 75 years old were eligible to participate in this study, since women belonging to this cohort are eligible for breast cancer screening in the Netherlands (RIVM, 2020). As part of this study’s pre-registration (https://osf.io/gk4rz), a power analysis using G*Power (2015) was performed. The power analysis showed that 158 participants were needed to find medium-sized effects (d = 0.25) with power set at 0.80. Based on this outcome, the aim was to find a minimum of 158 participants.

Participants were recruited via social media platforms, including Facebook and LinkedIn, by sharing a brief and concise introduction of the study together with the link to the survey (presented in Appendix A). Women between 50 and 75 years old were requested to fill out the survey. Furthermore, eligible women from the researcher’s personal networks were personally approached and were sent the survey directly. Participants were asked to send the survey to other eligible women in their circle, ultimately leading to a snowball effect.

Material

Scenario

A hypothetical scenario was used in which ‘risk level’ and ‘presentation format’ were manipulated. In this scenario, participants were told that they had participated in a breast cancer risk assessment of which they would now receive the results. These results were presented in a letter, containing a fictive personal risk estimate (i.e., either 7% or 28%) and the comparative risk estimate (i.e., 14%). The average, comparative risk estimate was 14%, since the average risk of a Dutch woman getting breast cancer in her lifetime is 1 out of 7, corresponding to 14% (RIVM, 2020). The above average (28%) and below average (7%) risk estimates were calculated by either multiplying or dividing the average risk estimate by two, respectively. These risk estimations were, depending on the assigned condition, presented numerical-only, numerical and in a bar graph, or numerical and in an icon array. To rule out possible effects of numerical format, it was chosen to present the numerical risks both in probability (e.g., 28%) and in frequency format (e.g., 28 out of 100 women). Thus, the structure of the letter was constant between conditions, while 'risk level' and 'presentation format' differed between conditions. The letter in the 28% risk level plus numerical-only condition and its English translation are presented in Table 1. The other five scenarios can be found in Appendix B.

Table 1

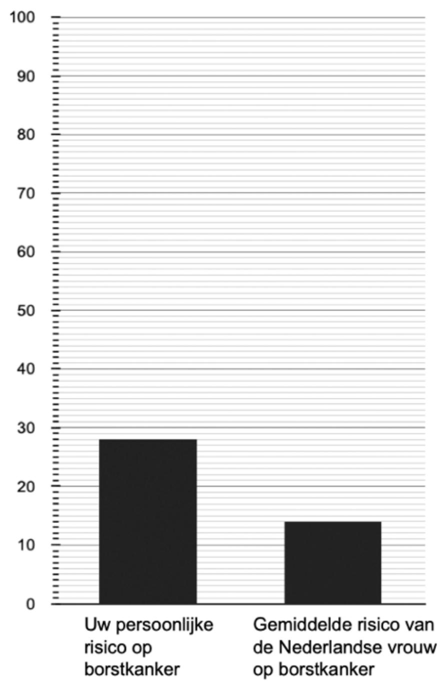
The Letter in the 28% plus Numerical-Only Condition and its English Translation

Dutch	English
Middels deze brief informeren wij u over de uitslag van uw risicobepaling.	Through this letter we are informing you about the result of your risk assessment.
Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.	Based on factors such as age, gender, age at menarche, age at first pregnancy and family history, your personal risk of getting breast cancer in your lifetime has been estimated.
Hieruit is gebleken dat uw risico op borstkanker 28% (28 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).	This has shown that your risk of getting breast cancer is 28% (28 out of 100 women). The mean risk of getting breast cancer for Dutch women is 14% (14 out of 100 women).

Both visualizations, bar graph and icon array, were created manually. They contained the same risk information, but differed in visual format. Both graphs were deliberately shown in black-and-white instead of color, since color could affect interpretations (Meyers-Levy & Peracchio, 1995; Stewart et al., 2009), which would be an additional and unintentional manipulation. Furthermore, the bar graph's y-axis was chosen to range from 0 to 100, since shortening the y-axis could be misleading and lead to less accurate interpretations (Yang et al., 2021). The bar graph and icon array format in the high-risk condition are depicted in Figure 2 and 3, respectively.

Figure 2

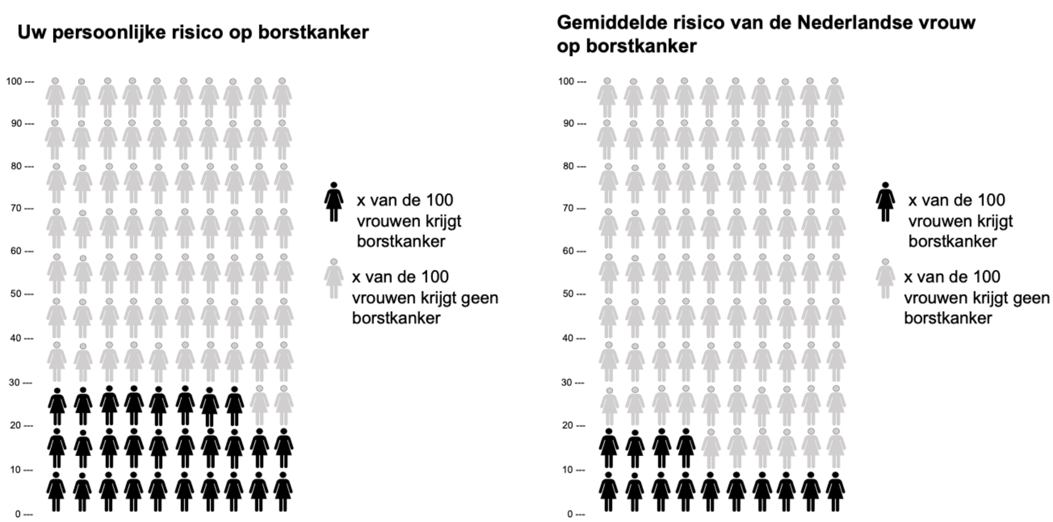
Bar Graph in the 28% Risk Condition



Note. Translation of the Dutch text: 'Your personal breast cancer risk', 'Average Dutch woman's breast cancer risk'.

Figure 3

Icon Array in the 28% Risk Condition



Note. Translation of Dutch titles: ‘Your personal breast cancer risk’, ‘Average Dutch woman’s breast cancer risk’. Translation of text next to icon array: ‘x out of 100 women get breast cancer’, ‘x out of 100 women do not get breast cancer’.

Outcome Measures

Manipulation Check. To check whether participants a) understood the numbers that were presented correctly, and b) could place themselves in the scenario, two manipulation checks were performed. The first manipulation check was: “How high is your own risk of developing breast cancer in comparison with the Dutch women’s average risk in this letter?”, with the answer options *below average, average, above average* or *I don’t know*. The second manipulation check included the item: “How easy or difficult did you find it to imagine that the breast cancer risk in this letter is applicable to you?”, whereby answers were given on a 5-point Likert scale ranging from *very easy* (1) to *very hard* (5).

Accuracy of Perceived Risk. ‘Accuracy of perceived risk’ was measured via an open-ended question that was taken from Schapira et al. (2004) and translated into Dutch: “What do you think your personal risk is of getting breast cancer in your lifetime?”. Answers were given on a scale from 0% (*no chance*) to 100% (*completely certain*). ‘Accuracy of perceived risk’ was determined by subtracting the given risk level - either 7% or 28% - from participants’ estimated risk, whereby scores closer to zero were deemed more accurate. The open-ended ‘perceived risk’ item can be found in Appendix C.

Perceived Risk. ‘Perceived risk’ was measured via two closed-ended questions: “How big or small do you think your risk of getting breast cancer in your life is?” and “How likely do you think it is that you will get breast cancer in your life?”. Both were answered on 5-point Likert scale, ranging from *very small* (1) to *very big* (5) and from *very unlikely* (1) to *very likely* (5), respectively. ‘Perceived risk’ scores were calculated by computing the mean score of these two items. The scale proved to be highly reliable at baseline level (α = .81). The ‘perceived risk’ items are presented in Appendix C.

Breast Cancer Worry. ‘Breast cancer worry’ was measured with three items that were taken from Lipkus et al. (2005) and translated into Dutch. An example of an item was: “How worried are you of getting breast cancer in your lifetime?”. Answers were given on a 6-point Likert scale with the end points labeled as *not at all* (1) and *extremely* (6). ‘Breast cancer worry’ scores were calculated by computing the mean score of these three items. The scale proved to be highly reliable at baseline level (α = .96). The items can be found in Appendix C.

Numeracy. ‘Numeracy’ was assessed with a Dutch version of the Subjective Numeracy Scale (SNS, Fagerlin et al., 2007b), a validated measure of quantitative ability and preference for receiving numerical information. The Dutch version was requested and adopted from Vromans et al. (2020). The SNS correlates strongly with objective numeracy and with ability to recall and comprehend both textual and graphical information (Zikmund-Fisher et al., 2007). The SNS score of each participant was calculated by computing the mean of the scores on the eight items, ranging from *least numerate* (1) to *most numerate* (6). The scale proved to be highly reliable (α = .87). The eight-item SNS is presented in Appendix C.

Procedure

Ethical approval from the Research Ethics and Data Management Committee of Tilburg University (REDC 2019.26c) was obtained before data collection. Data were gathered in May 2021 via a Qualtrics survey, which was written in Dutch. After clicking on the link to the survey, participants saw an information letter with explanations of the study and of the data and privacy guidelines. In order to move on to the actual survey, participants had to agree with these guidelines by providing informed consent. After informed consent was obtained, participants were shown demographic questions and were asked whether they had (had) breast cancer or were already familiar with their personal risk of getting breast cancer. Then, participants moved on to the items measuring ‘perceived risk’ and ‘breast cancer worry’. These were asked both before and after exposure to the hypothetical risk scenario in order to obtain baseline measurements. After that, participants were randomly shown one of six hypothetical risk scenarios. They were instructed to read these scenarios carefully and imagine as if the scenarios were applicable to them. Next, ‘perceived risk’ and ‘breast cancer worry’ items were presented again, in order to check how perceived risk and worry changed with respect to the first measurement. Lastly, participants filled out the SNS. After answering all questions, a debriefing was shown in which participants were thanked for their participation and the aim of the study was explained. On average, it took participants 9.95 minutes (*SD* = 12.19) to complete the survey.

Data Analysis

To test if participants within the six conditions differed in their mean age, subjective numeracy skills, second manipulation check and baseline risk perceptions and breast cancer worry, six one-way Analyses of Variance (ANOVAs) were performed. Moreover, three chi-square tests were performed to check if there were any differences between the six conditions in terms of mean education level, number of relatives with breast cancer and first manipulation check. Assumption evaluations were performed for each analysis separately, of which the results can be found in Appendix D.

To test whether accuracy of perceived risk, perceived risk and breast cancer worry differed by presentation format and risk level while controlling for numeracy, a two-way Multivariate Analysis of Covariance (MANCOVA) was performed. Independent variables were ‘presentation format’ and ‘risk level’. Dependent variables were ‘accuracy of perceived risk’ (i.e., estimated risk minus actual risk), ‘perceived risk’ and ‘breast cancer worry’ scores of the second measurement. The mean SNS score was included as a covariate. In case the MANCOVA showed a main effect of ‘presentation format’, an interaction effect, or both, planned contrast analyses (e.g., Helmert's, to compare both visual conditions and the numerical condition) or simple effects analyses were performed.

Finally, as an exploratory analysis, a three-way Multivariate Analyses of Variance (MANOVA) with 'presentation format', 'risk level' and 'numeracy height' as independent variables, and 'accuracy of perceived risk', 'perceived risk' and 'breast cancer worry' as dependent variables, was performed. In case the MANOVA showed a significant main effect of 'presentation format', any interaction effect(s), or both, planned contrast analyses or simple effects analyses were performed. It was chosen to base the distinction between lower and higher numeracy (i.e., 'numeracy height') on the median of the mean SNS scores (median = 4.25), since only a very small subset (*n* = 47) scored 3 or lower; the boundary for low and high numeracy (Zikmund-Fisher et al., 2014a). Women who scored below or on the median are hereafter referred to as less numerate (*n* = 207), whereas women who scored above the median are referred to as highly numerate (*n* = 196).

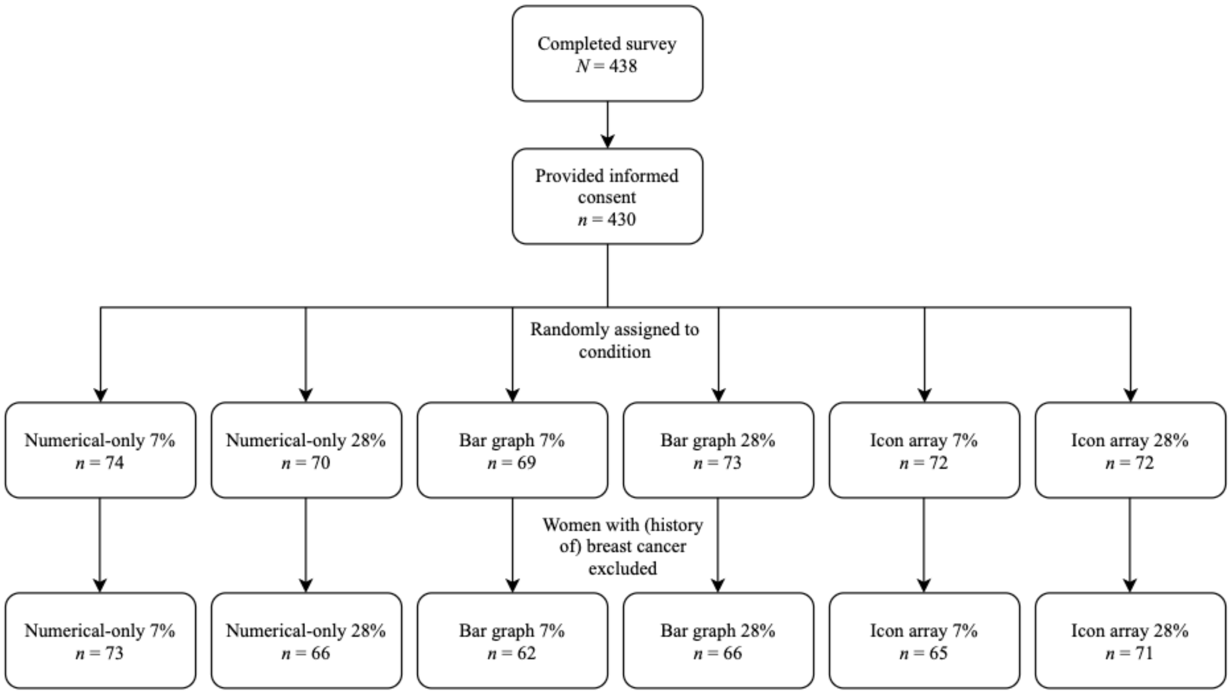
Results

Participants

A total of 438 participants completed the survey, of whom 430 provided informed consent. Data from 27 women were excluded because they indicated to have (had) breast cancer, leading to an ultimate number of 403 participants. A flowchart of the data collection and exclusion procedures is depicted in Figure 4.

Figure 4

Flowchart of the Data Collection and Exclusion Procedures



All participants were Dutch women between 50 and 75 years old (*M* = 58.44, *SD* = 6.37). Demographic characteristics of the 403 participants are presented in Table 2.

Table 2

Demographic Characteristics of the Participants per Condition

	Numerical-only		Bar graph		Icon array		
	7%	28%	7%	28%	7%	28%	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Age	58.40 (6.37)	57.85 (5.63)	58.58 (7.02)	60.05 (7.00)	57.86 (6.32)	57.93 (5.77)	.328
Numeracy	4.27 (0.91)	4.21 (0.98)	4.12 (0.90)	4.21 (0.83)	4.33 (0.89)	4.27 (0.99)	.856
	n (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	
Education							
Low[a]	6 (8.2%)	5 (7.6%)	8 (12.9%)	7 (10.6%)	3 (4.6%)	3 (4.2%)	
Intermediate[b]	24 (32.9%)	19 (28.8%)	20 (32.3%)	15 (22.7%)	13 (20.0%)	20 (28.2%)	
High[c]	43 (58.9%)	42 (63.6%)	34 (54.8%)	44 (66.7%)	49 (75.4%)	48 (67.6%)	.400
Relatives with breast cancer							

First degree	11 (15.1%)	6 (9.1%)	5 (8.1%)	11 (16.7%)	6 (9.2%)	10 (14.1%)	
Second degree	17 (23.3%	17 (25.8%)	15 (24.2%)	11 (16.7%)	12 (18.5%)	16 (22.5%)	
First and second degree	3 (4.1%)	4 (6.1%)	2 (3.2%)	2 (3.0%)	3 (4.6%)	3 (4.2%)	
None	42 (57.5%)	39 (59.1%)	40 (64.5%)	42 (63.6%)	44 (67.7%)	42 (59.2%)	.949

[a] Lower education is defined as elementary school (bo), lower vocational education (lbo) or lower general secondary education (mavo, mulo).

[b] Intermediate education is defined as secondary vocational education (mbo) or higher general secondary education (havo, vwo, atheneum).

[c] Higher education is defined as higher vocational education (hbo) or university (wo).

Manipulation Check

The first manipulation check showed that the majority of participants correctly interpreted their personal risk to be below or above average (70.7%). The second manipulation check showed that, on average, participants found it moderately difficult (M = 2.83, SD = 1.54) to place themselves in the scenario. Results of the manipulation check per condition are shown in Table 3.

Table 3

Results of Manipulation Check per Condition

	Numerical-only		Bar graph		Icon array		
	7%	28%	7%	28%	7%	28%	*p*
Correct interpretation, *n* (%)	52 (71.2%)	43 (65.2%)	49 (79.0%)	48 (72.7%)	48 (73.8%)	45 (63.4%)	.375
Difficulty placing themselves in scenario[a], *M* (*SD*)	2.12 (0.83)	2.48 (0.88)	1.98 (0.84)	2.62 (0.89)	2.08 (0.89)	2.59 (0.87)	.000

[a] Measured on a 5-point scale.

As can be seen in Table 3, participants did not differ in their interpretations. Regarding the second manipulation check, however, women in the 28% condition seemed to have more difficulty placing themselves in the scenario than women in the 7% condition. This was found in all three presentation format conditions. More detailed results can be found in Appendix D.

Effects of Presentation Format and Risk Level

The baseline and second measurement scores for (accuracy) of perceived risk and breast cancer worry are depicted in Table 4. As can be seen, none of the baseline measurement scores differed significantly between the six conditions, which means that the six conditions are comparable in terms of baseline risk perceptions and breast cancer worry.

Table 4

Means and Standard Deviations for (Accuracy of) Perceived Risk and Breast Cancer Worry per Condition

Baseline measurement	Numerical-only		Bar graph		Icon array		
	7%	28%	7%	28%	7%	28%	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*p*
Perceived risk (%)	30.16 (18.07)	30.00 (17.54)	34.71 (22.03)	35.39 (20.26)	31.51 (17.26)	36.37 (19.20)	.196
Perceived risk[a]	2.61 (0.62)	2.66 (0.61)	2.60 (0.64)	2.66 (0.65)	2.70 (0.49)	2.78 (0.57)	.531
Breast cancer worry[b]	2.46 (0.99)	2.63 (0.96)	2.56 (1.12)	2.51 (1.13)	2.43 (0.92)	2.85 (1.13)	.180
Second measurement	**Numerical-only**		**Bar graph**		**Icon array**		
	7%	28%	7%	28%	7%	28%	
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	
Perceived risk (%)	19.21 (19.09)	29.92 (16.02)	17.32 (15.87)	34.12 (17.34)	15.80 (15.70)	32.96 (16.33)	

Accuracy of perceived risk[c]	12.21 (19.09)*	1.92 (16.02)*	10.32 (15.87)*	6.12 (17.34)*	8.80 (15.70)*	4.96 (16.33)*
Perceived risk[a]	2.31 (0.75)*	2.90 (0.67)*	2.24 (0.67)*	2.92 (0.65)*	2.28 (0.62)*	2.96 (0.69)*
Breast cancer worry[b]	2.29 (0.99)*	3.08 (1.10)*	2.20 (0.89)*	2.94 (1.14)*	2.08 (0.87)*	3.28 (1.05)*

* = significant difference (p < .001) by risk level.

[a] Measured on a 5-point scale.

[b] Measured on a 6-point scale.

[c] Scores closer to zero are more accurate.

Assumption evaluations were performed. Results from these evaluations can be found in Appendix D. Since not all assumptions were met, more weight is placed on the Pillai’s Trace statistic, because this statistic is more robust against assumption violations.

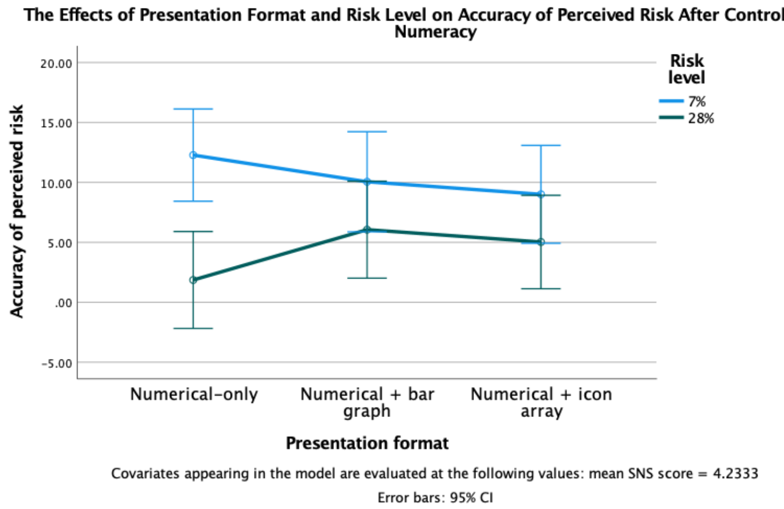
With the use of Pillai’s Trace criterion, the combined dependent variables were significantly different by risk level (V = .34, F(3, 394) = 68.81, p < .001, partial η^2 = .34) but not by presentation format (V = .01, F(6, 790) = 0.40, p = .882), after controlling for numeracy. Also, no significant interaction was found (V = .02, F(6, 790) = 1.54, p = .161) after controlling for numeracy. Lastly, the effect of numeracy was marginally significant (V = .02, F(3, 394) = 2.58, p = .053, partial η^2 = .02).

Risk level had a main effect on accuracy of perceived risk (F(1, 396) = 13.53, p < .001, partial η^2 = .03), perceived risk (F(1, 396) = 93.30, p < .001, partial η^2 = .19), and breast cancer worry (F(1, 396) = 80.75, p < .001, partial η^2 = .17). Participants had more accurate risk perceptions, higher perceived risk and more breast cancer worry in the above average (28%) condition than in the below average (7%) condition. Thus, the first hypothesis can be confirmed: Risk perceptions and breast cancer worry were indeed higher in the above average condition than in the below average condition. The second hypothesis, however, cannot be confirmed: The effect of risk level was not strengthened by presentation format.

Presentation format did not have an effect on accuracy of perceived risk ($F(2, 396) = 0.16$, $p = .852$), perceived risk ($F(2, 396) = 0.18$, $p = .838$) and breast cancer worry ($F(2, 396) = 0.51$, $p = .600$). Therefore, the third hypothesis cannot be confirmed: Numerical plus visual format does not lead to higher accuracy of perceived risk than numerical-only format, and icon arrays do not lead to the highest accuracy. The effects of risk level and presentation format on the three outcome measures are presented in Figures 5, 6 and 7.

Figure 5

Relationship Between Presentation Format, Risk Level and Accuracy of Perceived Risk



Note. Scores closer to zero are more accurate.

Figure 6

Relationship Between Presentation Format, Risk Level and Perceived Risk

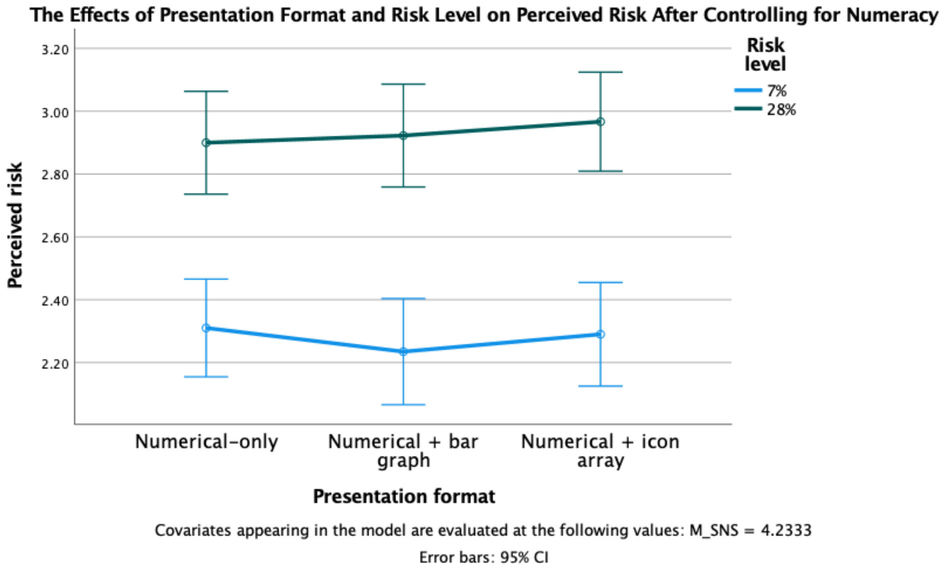
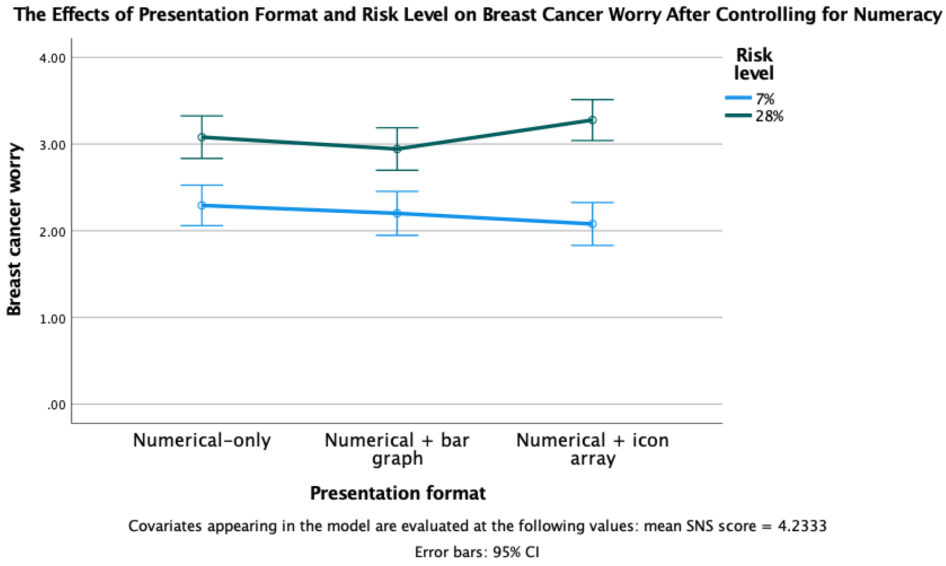


Figure 7

Relationship Between Presentation Format, Risk Level and Breast Cancer Worry



Exploratory Analysis: Less Versus Highly Numerate

Assumption evaluations were performed. Results from these evaluations can be found in Appendix D. Since not all assumptions were met, more weight is placed on the Pillai's Trace statistic, because this statistic is more robust against assumption violations. Table 5 provides the most important results of the analysis. A complete overview of the statistics for less and highly numerate women is presented in Appendix E.

Table 5

Means and Standard Deviations for (Accuracy of) Perceived Risk and Breast Cancer Worry per Risk Level, Divided by Numeracy

	Less numerate (score 0 – 4.25)		Highly numerate (score 4.26 – 6)	
	7%	28%	7%	28%
	n = 97	*n* = 110	*n* = 103	*n* = 93
Perceived risk (%)	20.76 (18.55)	32.06 (18.11)	14.46 (14.94)	32.69 (14.65)
Accuracy of perceived risk[a]	13.76 (18.55)**	4.06 (18.11)	7.46 (14.94)**	4.69 (14.65)
Perceived risk[b]	2.45 (0.65)**	2.88 (0.73)	2.12 (0.67)**	2.99 (0.59)
Breast cancer worry[c]	2.34 (1.00)	2.97 (1.10)*	2.06 (0.82)	3.26 (1.08)*

* = significant difference ($p < .05$) between less and highly numerates.

** = significant difference ($p < .01$) between less and highly numerates.

[a] Scores closer to zero are more accurate.

[b] Measured on a 5-point scale.

[c] Measured on a 6-point scale.

The MANOVA showed no main effect of presentation format ($V = .01$, $F(6, 780) = 0.42$, $p = .864$), no main effect of numeracy ($V = .02$, $F(3, 389) = 2.00$, $p = .114$), no interaction effect between presentation format and risk level ($V = .02$, $F(6, 780) = 1.43$, $p = .201$), no interaction effect between presentation format and numeracy ($V = .01$, $F(6, 780) =$

0.43, p = .862), and no interaction effect between presentation format, risk level and numeracy (V = .00, F(6, 780) = 0.16, p = .987). However, the MANOVA showed a significant main effect of risk level (V = .35, F(3, 389) = 68.20, p < .001, partial η^2 = .35) and a significant interaction effect between risk level and numeracy (V = .03, F(3, 389) = 3.81, p = .010, partial η^2 = .03).

Risk level had a main effect on perceived risk (F(1, 391) = 92.42, p < .001, partial η^2 = .19), breast cancer worry (F(1, 391) = 81.51, p < .001, partial η^2 = .17) and accuracy of perceived risk (F(1, 391) = 13.72, p < .001, partial η^2 = .03). Women had higher risk perceptions, more accurate risk perceptions, and more breast cancer worry in the above average (28%) risk condition than in the below average (7%) risk condition. This finding is similar to the results of the main analysis.

The interaction effect between risk level and numeracy was present for perceived risk (F(1, 391) = 10.35, p = .001, partial η^2 = .03), breast cancer worry (F(1, 391) = 7.19, p = .008, partial η^2 = .02) and accuracy of perceived risk (F(1, 391) = 4.34, p = .038, partial η^2 = .01). Follow-up simple effects analyses revealed that less and highly numerate women differed on perceived risk (F(1, 399) = 12.09, p = .001) and accuracy of perceived risk (F(1, 399) = 7.12, p = .008) in the below average (7%) risk condition, such that less numerate women had higher risk perceptions and less accurate risk perceptions in the below average condition than highly numerate women. Regarding breast cancer worry, there was a marginally significant difference between less and highly numerate women in the below average (7%) condition (F(1, 399) = 3.73, p = .054) and a significant difference in the above average (28%) condition (F(1, 399) = 4.15, p = .042). This means that less numerate women had slightly more breast cancer worry in the below average condition and significantly less worry in the above average condition than highly numerate women. Visual representations of these interactions are presented in Appendix E.

Discussion

The aim of this study was to examine the effects of risk level and presentation format on both cognitive (i.e., risk perceptions) and affective (i.e., breast cancer worry) responses to comparative breast cancer risk information. This was not yet investigated in previous research, while it is crucial knowledge for medical professionals or designers of information leaflets and letters in the context of personalized breast cancer screening. In this study, Dutch women who are eligible for breast cancer screening were shown a hypothetical scenario containing fictive risk estimates. These risk estimates were either below (7%) or above (28%) average (i.e., risk level), and were presented numerical-only, numerical and in a bar graph, or numerical and in an icon array (i.e., presentation format). The effects of risk level and presentation format on accuracy of perceived risk, perceived risk and breast cancer worry were investigated.

Findings

The first hypothesis was that above average (i.e., 28%) risk estimates lead to higher perceived risk and breast cancer worry than below average (i.e., 7%) risk estimates. This hypothesis was confirmed: A 28% personal risk estimate indeed led to higher perceived risk and breast cancer worry than a 7% estimate. The second hypothesis, which stated that presentation format would strengthen the effect of risk level such that risk perceptions and breast cancer worry were higher in the numerical plus icon array condition than in the other two conditions, cannot be confirmed. Presentation format did not interact with risk level. In sum, risk level significantly affected perceived risk and breast cancer worry, but presentation format did not. The latter was also not the case when comparing less versus highly numerate women. However, less numerate women had higher risk perceptions than highly numerate women in the below average condition and were less worried in the above average condition than highly numerate women, indicating that numeracy influences the effects of risk level on perceived risk and breast cancer worry.

In contrast to the third hypothesis, the results did not show an effect of presentation format on accuracy of perceived risk. Women did not have more accurate risk perceptions when given their personal and comparative risk estimate both numerically and visually: Numerical plus icon array format did not lead to higher accuracy than numerical plus bar graph or numerical-only format. Although no *a priori* hypotheses were formulated about the effect of risk level on accuracy of perceived risk, the results revealed a significant effect of risk level. Providing an above average (28%) personal risk estimate led to more accurate risk perceptions than providing a below average (7%) risk estimate. Exploratory analyses revealed that numeracy interacted with risk level: Less numerate women were less accurate in the below average condition than highly numerate women, whereas there was no difference in the above average risk condition. This implicates that highly numerate women are less influenced by risk level. The effects of risk level on accuracy of perceived risk were unexpected but nonetheless important.

The results for hypothesis one are generally in accordance with previous research: An above average risk level leads to higher perceived risk and more breast cancer worry than a below average risk level. French et al. (2004) also found that unfavourable comparative risk information (i.e., higher personal risk than comparative risk) led to higher risk perceptions and more worry than favourable comparative risk information (i.e., lower personal risk than comparative risk). This was also the case in the present study. Therefore, the findings implicate that varying personal risk levels evoke responses similar to varying average risk levels, as illustrated by Fagerlin et al (2007a): Being below or above average influences how one perceives these risks. Important to note is that the effects of the 7% and 28% risks are assumed to be due to the fact that they are below and above average, respectively. It is, however, not ruled out that the effects are due to the numbers themselves, regardless of them being below or above average.

The finding that presentation format does not play a role regarding perceived risk and breast cancer worry, as was expected in the second hypothesis, is not in line with previous research. As demonstrated by Schapira et al. (2006) and Timmermans et al. (2008), risks of similar magnitude are perceived as higher and more worrisome in icon array format than in bar graph format or numerical-only format. The present study did not find such results. Only risk level affected risk perceptions and breast cancer worry, and this was not strengthened by presentation format. This could mean that presentation format does not evoke similar responses in the realm of comparative risk information: Women’s responses to comparative risk information might not be influenced by presentation format, whereas presentation format does seem to affect women’s responses to personal risk information.

The findings for the third hypothesis are also not in line with previous research. As mentioned by Zipkin et al. (2014) in their systematic review of 91 studies, eight studies found advantages of the addition of visual displays, whereas only two small studies did not find such advantages. Moreover, several studies on the presentation of personal risk information did find that visualizations, and icon arrays in particular, improve the accuracy of risk perceptions (see, e.g., Brown et al., 2011; Galesic et al., 2009). The present study did not find such advantages of the addition of visualizations. The finding that risk level affected accuracy of perceived risk, but presentation format did not, is an unexpected and novel finding, although this effect was smaller than the effects of risk level the other two outcome measures and the effect only seemed to be present for less numerate women. This indicates that less numerate women could have difficulty interpreting low(er than average) risks (Kreuzmair et al., 2016), which means that numeracy plays a role in the interpretation of risk information. This corresponds with previous research. For instance, Garcia-Retamero and Hoffrage (2013) and

Keller and Siegrist (2009) also found that one's risk perceptions and accuracy of interpretations are influenced by one's numerical skills, such that highly numerates are less affected by format than less numerates.

Multiple reasons could explain why presentation format did not directly influence accuracy of perceived risk, nor affected the relationship between risk level and perceived risk and breast cancer worry. A first possible reason has to do with the nature of comparative risk information. As highlighted by Zikmund-Fisher (2019), the function of comparative risk information is to provide context. People use comparisons when interpreting risk information (Fagerlin et al., 2007a; Klein, 1997). Since they rely on this comparison by interpreting the comparative, average risk estimate, presentation format could have become redundant. There was already enough context, so an additional visualization might have been unnecessary and was therefore not used by the participants. Moreover, numerical risk estimates were provided in all three presentation format conditions, which also could have made the visualizations redundant.

Another reason could be that women did read the scenario accurately, but did not believe it was applicable to them. In other words, that the participants did not trust the given fictive risk estimate. This is also shown by the results of the second manipulation check. Women indicated that they found it moderately difficult to place themselves in the scenario and to imagine that the fictive risk estimates were applicable to them. Moreover, the results revealed that women in the above average (28%) condition had significantly more difficulty placing themselves in the scenario than women in the below average (7%) condition. This is an unexpected and interesting finding, especially because women in the above average condition had more accurate risk perceptions. An explanation could be that the 28% fictive risk estimate was close to their self-estimated risk at baseline, which was 33.98% in the three 28% risk conditions. Women could have been surprised that the 28% risk estimate was so close to their self-estimated risk, which made them question if the fictive risk estimate was truly fictive or was somehow adjusted to their baseline estimate.

Study Limitations and Future Research

A first limitation is that fictive breast cancer risk estimates were used instead of real risk estimates, although hypothetical scenarios are more often used in risk communication research (see, e.g., Fagerlin et al., 2007a; Hawley et al., 2008; Schapira et al., 2006). Moreover, women overall indicated to not find it extremely difficult to place themselves in the scenario. It is, however, still recommended for future studies to include real risk estimates instead of fictive ones when possible, since this comes closest to the real-world situation. Regarding the generalizability of the findings, the size and characteristics of the participant sample were close to the natural situation: The sample size was large and all participants were Dutch women between 50 and 75 years old, which is the cohort that is eligible for breast cancer screening in the Netherlands (RIVM, 2020). Other studies oftentimes make use of student samples, which is not representative for the real population. However, participants in the present study were mainly highly educated and had relatively high numerical skills (i.e., median = 4.25). Therefore, future research could examine whether a larger group of lowly educated women, or a larger group of women with low numerical skills (i.e., SNS score lower than 3) (Zikmund-Fisher et al., 2014a), respond similarly or differently to variations in risk level and presentation format when provided comparative risks.

A second limitation is the way the survey was built up. This could have affected participants' answers on – at least – the open-ended perceived risk question. Participants had to estimate their baseline risk perception before being exposed to the scenario with the fictive risks, which could have affected their perception of the eventual fictive risks and subsequently their answer on the second risk perception estimate. This has been shown by Fagerlin et al. (2005). Fagerlin and colleagues (2005) asked women to estimate their lifetime breast cancer risk and found that many women tend to overestimate their risk (46%). When shown the actual risk estimate (13%), this actual risk feels relatively low compared to the much higher self-estimated risk, subsequently leading to relief. Therefore, it is possible that women's baseline self-estimation affected their response to the second self-estimation. This makes it difficult to distinguish whether the risk estimate at second measurement is a result of the presented scenario or of participants' risk estimate at baseline. The results demonstrate the possibility of this effect. The mean baseline lifetime breast cancer risk perception was 33%, which corresponds closely to the fictive 28% estimate half of these women received. Therefore, these women's risk perceptions could be more accurate, whereas this effect might be different when either their baseline risk perception or the given personal risk estimate differed. It is advisable for future studies to exclude the baseline open-ended perceived risk question in order to eliminate possible adverse effects.

The results of the present study leave room for future research. Since it was found that presentation format does not affect any of the dependent variables, it is possible that – at least concerning risk perceptions and breast cancer worry – visualizations become redundant when presenting comparative risk information, since context is already provided. Moreover, numerical risks were given in all three presentation format conditions, which also could have made the visualizations redundant. To test if people truly ignore visualizations, it is advisable to conduct eye tracking studies to investigate visual search strategies of people when interpreting risk information. A previous eye tracking study on the processing of icon arrays showed that highly numerates focus on concrete, numerical information by counting the highlighted icons, while less numerates relied more on holistic processing of the visual display (Kreuzmair et al., 2016). Furthermore, the finding that less numerate women are more accurate in the above average (28%) than in the below average (7%) condition also leaves room for future studies. More research is needed to see if the effect of risk level still holds when providing different risk estimates, such as real instead of fictive ones, or risks that are similar to or only slightly deviant from the average risk.

Lastly, it is possible that visualizations do not, but other presentation formats do matter for comparative risk information. For example, colour could affect one’s interpretation of risk information (e.g., red is negative, green is positive) (Meyers-Levy & Peracchio, 1995; Stewart et al., 2009) and tables could work well too (Hawley et al., 2008). Therefore, it is interesting to further investigate other presentation formats. Besides, it is possible that presentation format does matter, but just not for the outcomes measured in this study. Other studies, such as Tait et al.’s (2010), did find positive effects of visualizations in personal risk communication, but they focused on measures like verbatim and gist knowledge instead of accuracy of risk perceptions. Thus, future research could test whether visualizations in comparative risk communication do work better for verbatim and gist understanding of the information communicated. This is evidential, since comparative risk information is thought to evoke intuitive responses to risks (i.e., gist understanding), while personal risk information potentially only relates to the numerical presentation of a risk (i.e., verbatim understanding) (Lloyd et al., 2001; Reyna, 2004; Schwartz, 2009; Windschitl et al., 2002).

Implications

The present study implicates that the presentation format in which comparative risk information is provided does not affect cognitive nor affective responses. This means that designers of information tools, leaflets or letters may consider including visualizations, since it does not affect the interpretation of below or above average risks directly nor indirectly. This was also not the case for less numerate women. Since an extensive body of studies acknowledges the benefits of including visualizations (see, e.g., Garcia-Retamero & Cokely, 2017; Lipkus, 2007; Spiegelhalter, 2017; Zipkin et al., 2014), it is advisable to still include visualizations, be it bar graphs or icon arrays, because this does not negatively influence cognitive nor affective responses to comparative risk information.

Regarding the effects of risk level, it is advisable for designers of breast cancer risk letters, brochures or leaflets to take into account that the height of one's personal risk level affects the accuracy of the interpretation of that risk, and that numeracy affects this relation as well – at least for comparative risk information. Less numerate women have more accurate risk perceptions when their personal risk level is higher than average (28%), but highly numerate women are less affected by the height of their personal risk level. This could mean that when communicating low or lower than average risks to (less numerate) women, this should be done in a clear and effective way, such that their risk perceptions are as accurate as possible. It has been argued that 'less is more': Only the information that is needed should be provided, with the meaning of the most important information highlighted, subsequently reducing the cognitive effort needed (Peters et al., 2007a; Peters et al., 2007b).

This study contributes to the body of research on comparative risk information and communication in the sense that it provides novel insights into comparative breast cancer risk communication. It is, to my knowledge, the first study to examine the effects of presentation format on the provision of below and above average risks. The results implicate that the effects of presentation format are different for comparative risk information than for personal risk communication, since no effects of presentation format were found, whereas studies on personal risk communication commonly demonstrate advantages of visualizations (see, e.g., Garcia-Retamero & Hoffrage, 2013; Schapira et al., 2001; Waters et al., 2007). The present study opens room for future studies to gain deeper insight into the communication of comparative risk information.

Conclusion

In conclusion, there was no effect of presentation format on (accuracy of) perceived risk nor on breast cancer worry. There was, however, an effect of risk level on all three outcomes: Women had more accurate risk perceptions, higher risk perceptions, and more breast cancer worry in the above average (28%) condition than in the below average (7%) condition. This means that risk level does affect women's risk perceptions and breast cancer worry, but this effect is not strengthened by presentation format. Moreover, the exploratory analysis revealed that numeracy impacts the effects of risk level, such that risk perceptions and breast cancer worry were different for less and highly numerate women. Regarding accuracy of perceived risk, only less numerate women seemed to be less accurate in the below average condition, indicating that highly numerate women are less influenced by the height of their personal risk level. The present study contributes to the body of research on comparative risk information in the sense that it provides novel insights into the effects of visualizations and risk levels on cognitive and affective responses, with numeracy as a contributing factor.

What does this study mean for women who might be invited for personalized breast cancer screening? Imagine that you are the same 50-year old woman, but now a few days later. Today you receive the results of your breast cancer risk assessment. When opening the letter, you see both text and an icon array, representing your risk and the average risk of Dutch women. Your risk is 28%, while the average woman's risk is 14%. The results match your initial cognitions and feelings: Your risk is higher than average. The icon array shows you that, in your risk category, 27 other women out of 100 are at higher than average risk. Although you belong to the above average risk category, it makes you feel understood. You are not alone, and there are things that can be done with these results: Taking part in personalized breast cancer screening, in which your personal situation will be considered.

References

Aldrich, F. K., & Sheppard, L. (2000). Graphicacy: The fourth R?. *Primary Science Review, 64*, 8-11. http://tecfa.unige.ch/tecfa/teaching/LME/betrancourt/aldrich2000.pdf

Berry, D. (2004). *Risk, communication, and health psychology*. Open University Press.

Brown, S. M., Culver, J. O., Osann, K. E., MacDonald, D. J., Sand, S., Thornton, A. A., Grant, M., Bowen, D. J., Metcalfe, K. A., Burke, H. B., Robson, M. E., Friedman, S., & Weitzel, J. N. (2011). Health literacy, numeracy, and interpretation of graphical breast cancer risk estimates. *Patient Education and Counseling, 83*(1), 92-98. https://doi.org/10.1016/j.pec.2010.04.027

Brun, W. (1994). Risk perception: Main issues, approaches and findings. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 295-320). John Wiley and Sons.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210. https://doi.org/10.1007/bf01320076

Cleveland, W., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79*(387), 531–554. https://doi.org/10.2307/2288400

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*, 294-299. https://doi.org/10.1016/j.pec.2004.08.007

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2007a). “If I’m better than average, then I’m ok?”: Comparative information influences beliefs about risk and benefits. *Patient Education and Counseling, 69*(1-3), 140-144. https://doi.org/10.1016/j.pec.2007.08.008

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., Smith, D. M. (2007b). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale. *Medical Decision Making, 27*(5), 672-680. https://doi.org/10.1177/0272989X07304449

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140. https://doi.org/10.1177%2F001872675400700202

French, D. P., Sutton, S. R., Marteau, T. M., & Kinmoth, A. L. (2004). The impact of personal and social comparison information about health risk. *British Journal of Health Psychology, 9*(2), 187-200. https://doi.org/10.1348/135910704773891041

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28*(2), 210-216. https://doi.org/10.1037/a0014474

Garcia-Retamero, R., & Cokely, E. T. (2017). Designing visual aids that promote risk literacy: A systematic review of health research and evidence-based design heuristics. *Human Factors, 59*(4), 582-627. https://doi.org/10.1177/0018720817690634

Garcia-Retamero, R., & Galesic, M. (2010). Who profits from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine, 70*(7), 1019 -1025. https://doi.org/10.1016/j.socscimed.2009.11.031

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine, 83*, 27-33. https://doi.org/10.1016/j.socscimed.2013.01.034

Ghosh, K., Crawford, B. J., Pruthil, S., Williams, C. I., Neal, L., Sandhu, N. P., Johnson, R. E., Wahner-Roedler, D., Britain, M. K., Cha, S. S., & Ghohs, A. K. (2008). Frequency format diagram and probability chart for breast cancer risk communication: A prospective, randomized trial. *BMC Women's Health, 8*(18), 1-8. https://doi.org/10.1186/1472-6874-8-18

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine, 29*(4), 375-376. https://doi.org/10.1016/j.amepre.2005.06.012

Hawley, S. T., Zikmund-Fisher, B. J., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling, 73*(3), 448-455. https://doi.org/10.1016/j.pec.2008.07.023

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73*(5), 538-540. https://doi.org/10.1097/00001888-199805000-00024

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*(5500), 2261-2262. https://doi.org/10.1126/science.290.5500.2261

Keller, C., & Siegrist, M. (2009). Effect of risk communication formats on risk perception depending on numeracy. *Medical Decision Making, 29*, 483-490. https://doi.org/10.1177/0272989X09333122

Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology, 72*(4), 763–774. https://doi.org/10.1037//0022-3514.72.4.763

Kreuter, M. W., Lezin, N. A., & Kreuter, M. W., Green, L. W. (1997). *Community health promotion ideas that work*. Jones and Bartlett.

Kreuter, M. W., Strecher, V. J., & Glassman, B. (1999). One size does not fit all: The case for tailoring print materials. *Annals of Behavioral Medicine, 21*(4), 276–283. https://doi.org/10.1007/BF02895958

Kreuzmair, C., Siegrist, M., & Keller, C. (2016). High numerates count icons and low numerates process large areas in pictographs: Results of an eye-tracking study. *Risk Analysis, 36*(8), 1599-1614. https://doi.org/10.1111/risa.12531

Lichtenstein, S., & Slovic, P. (2006). *The construction of preference*. Cambridge University Press. https://doi.org/10.1017/CBO9780511618031

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making, 27*, 696-713. https://doi.org/10.1177/0272989X07307271

Lipkus, I. M., & Hollands, J. G. (1999). The visual communication of risk. *JNCI Monographs, 25*(25), 149-163. https://doi.org/10.1093/oxfordjournals .jncimonographs.a024191

Lipkus, I. M., Klein, W. M. P., Skinner, C. S., & Rimer, B. K. (2005). Breast cancer risk perceptions and breast cancer worry: What predicts what? *Journal of Risk Research, 8*(5), 439-452. https://doi.org/10.1080/1366987042000311018

Lloyd, A., Hayes, P., Bell, P. R. F., & Naylor, A. R. (2001). The role of risk and benefit perception in informed consent for surgery. *Medical Decision Making, 21*(2), 141-149. https://doi.org/10.1177%2F0272989X0102100207

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389-401. https://doi.org/10.1037/0022-0663.86.3.389

McCaul, K. D., Canevello, A. B., Mathwig, J. L., & Klein, W. M. P. (2003) Risk communication and worry about breast cancer. *Psychology, Health, and Medicine, 8*(4), 379–389. https://doi.org/10.1080/13548500310001604513

Meyers-Levy, J., & Peracchio, L. A. (1995). Understanding the effects of color: How the correspondence between available and required resources affects attitudes. *Journal of Consumer Research, 22*(2), 121-138. https://doi.org/10.1086/209440

Onega, T., Beaber, E. F., Sprague, B. L., Barlow, W. E., Haas, J. S., Tosteson, A. N. A., Schnall, M. D., Armstrong, K., Schapira, M. M., Geller, B., Weaver, D. L., & Conant, E. F. (2014). Breast cancer screening in an era of personalized regimens: A conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. *Cancer, 120*(19), 2955-2964. https://doi.org/10.1002/cncr.28771

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

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007b). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs, 26*(3), 741-748. https://doi.org/10.1377/hlthaff.26.3.741

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407-413. https://doi.org/10.1111/j.1467-9280.2006.01720.x

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-process approach. *Current Directions in Psychological Science, 13*(2), 60-66. https://doi.org/10.1111/j.0963-7214.2004.00275.x

Reyna, V. F., & Brainerd C. J. (2008) Numeracy, ratio bias and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 18*(1), 89-107. https://doi.org/10.1016/j.lindif.2007.03.011

Rijksinstituut voor Volksgezondheid en Milieu. (2020). *Factsheet bevolkingsonderzoek borstkanker*. Retrieved from https://www.rivm.nl/documenten/factsheet-borstkanker

Román, M., Sala, M., Domingo, L., Posso, M., Louro, J., & Castells, X. (2019). Personalized breast cancer screening strategies: A systematic review and quality assessment. *PLoS ONE, 14*(12), 1-18. https://doi.org/10.1371/journal.pone.0226352

Schapira, M. M., Davids, S. L., McAuliffe, T. L., & Nattinger, A. B. (2004). Agreement between scales in the measurement of breast cancer risk perceptions. *Risk Analysis, 24*(3), 665-673. https://doi.org/10.1111/j.0272-4332.2004.00466.x

Schapira, M. M., Nattinger, A. B., & McAuliffe, T. L. (2006). The influence of graphic format on breast cancer risk communication. *Journal of Health Communication, 11*(6), 569-582. https://doi.org/10.1080/10810730600829916

Schapira, M. M., Nattinger, A. B., & McHorney, C. A. (2001). Frequency or probability? A qualitative study of risk communication formats used in health care. *Medical Decision Making, 21*(6), 459-467. https://doi.org/10.1177/0272989X0102100604

Schwartz, P. H. (2009). Disclosure and rationality: Comparative risk information and decision-making about prevention. *Theoretical Medicine and Bioethics, 30*, 199-213. https://doi.org/10.1007/s11017-009-9111-7

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*(11), 966-972. https://doi.org/10.7326/0003-4819-127-11-199712010-00003

Slovic, P. (1999). Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. *Risk Analytics, 19*(4), 689-701. https://doi.org/10.1023/a:1007041821623

Spiegelhalter, D. (2017). Risk and uncertainty communication. *Annual Review of Statistics and its Application, 4*(1), 31-60. https://doi.org/10.1146/annurev-statistics-010814-020148

Stewart, B. M., Cipolla, J. M., & Best, L. A. (2009). Extraneous information and graph comprehension: Implications for effective design choices. *Campus-Wide Information Systems, 26*(3), 191-200. https://doi.org/10.1108/10650740910967375

Suls, J., & Wheeler, L. (2000). *Handbook of social comparison: Theory and research.* Kluwer Academic.

Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010). The effect of format on parents' understanding of the risks and benefits of clinical research: A comparison between text, tables, and graphics. *Journal of Health Communication, 15*(5), 487-501. https://doi.org/10.1080/10810730.2010.492560

Thompson, P. B., & Dean, W. (1996). Competing conceptions of risk. *Risk, Health, Safety, and Environment, 7*(4), 361-384.

Timmermans, D. R. M., Ockhuysen-Vermey, C. F., & Henneman, L. (2008). Presenting health risk information in different formats: The effect on participants' cognitive and emotional evaluation and decisions. *Patient Education and Counseling, 73*(3), 443-447. https://doi.org/10.1016/j.pec.2008.07.013

Vromans, R. D., Pauws, S. C., Bol, N., van de Poll-Franse, L. V., & Krahmer, E. J. (2020). Communicating tailored risk information of cancer treatment side effects: Only words or also numbers? *BMC Medical Informatics and Decision Making, 20*(277), 1-12. https://doi.org/10.1186/s12911-020-01296-7

Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. M. (2007). Reducing aversion to side effects in preventive medical treatment decisions. *Journal of Experimental Psychology: Applied, 13*(1), 11-21. https://doi.org/10.1037/1076-898X.13.1.11

Windschitl, P. D., Martin, R., & Flugstad, A. R. (2002). Context and the interpretation of likelihood information: The role of intergroup comparisons on perceived vulnerability. *Journal of Personality and Social Psychology, 82*(5), 742- 755. https://doi.org/10.1037/0022-3514.82.5.742

Wood, M. E., Farina, N. H., Ahern, T. P., Cuke, M. E., Stein, J. L., Stein, G. S., & Lian, J. B. (2019). Towards a more precise and individualized assessment of breast cancer risk. *Aging, 11*(4), 1305-1316. https://doi.org/10.18632%2Faging.101803

Yang, B. W., Vargas-Restrepo, C., Stanley, M. L., & Marsh, E. J. (2021). Truncating bar graphs persistently misleads viewers. *Journal of Applied Memory and Cognition*. https://doi.org/10.1016/j.jarmac.2020.10.002

Yates, J., & Stone, E. (1992). The risk construct. In J. Yates (Ed.), *Risk-taking behavior* (pp. 1-25). Wiley.

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*(1), 37S-49S. https://doi.org/10.1177/1077558712458541

Zikmund-Fisher, B. J. (2019). Helping people know whether measurements have good or bad implications: Increasing the evaluability of health and science data communications. *Policy Insights from the Behavioral and Brain Sciences, 6*(1), 29-37. https://doi.org/10.1177/2372732218813377

Zikmund-Fisher, B. J., Exe, N. L., & Witteman, H. O. (2014a). Numeracy and literacy independently predict patients’ ability to identify out-of-range results. *Journal of Medical Internet Research, 16*(8), e187. https://doi.org/10.2196/jmir.3241

Zikmund-Fisher, B. J., Fagerlin, A., &, Ubel, P. A. (2004). “Is 28% good or bad?” Evaluability and preference reversals in health care decisions. *Medical Decision Making*, *24*, 142-148. https://doi.org/10.1016/j.pec.2007.08.008

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making*, *27*(5), 663–671.

Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., Valerio, M., Holtzman, L. G., Scherer, L. D., & Fagerlin, A. (2014b). Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Medical Decision Making*, *34*(4), 443-453. https://doi.org/10.1177/0272989X13511706

Zipkin, D. A., Umscheid, C. A., Keating, N. L., Allen, E., Aung, K., Beyth, R., Kaatz, S., Mann, D. M., Sussman, J. B., Korenstein, D., Schardt, C., Nagi, A., Sloane, R., & Feldstein, R. A. (2014). Evidence-based risk communication: A systematic review. *Annals of Internal Medicine*, *161*(4), 270-280. https://doi.org/10.7326/M14-0295

Appendix A: Recruitment Post for Social Media

Voor mijn masterscriptie ben ik op zoek naar vrouwen tussen de 50 en 75 jaar die een korte vragenlijst (7-10 minuten) in willen vullen. Het zou geweldig zijn als u me daarbij kunt helpen!

De vragenlijst gaat over een mogelijk nieuwe aanpak van het bevolkingsonderzoek borstkanker. Hierbij wordt het bevolkingsonderzoek 'op maat' afgestemd op iemands persoonlijke risico op het krijgen van borstkanker. In mijn scriptieonderzoek bestudeer ik de interpretatie van deze risico's. Geen zorgen, uw persoonlijke risico wordt niet bepaald: dat wat u leest is fictief.

De vragenlijst is hier te vinden: https://lnkd.in/gAYrq2G

Behoort u zelf niet tot de doelgroep, maar wilt u toch een steentje bijdragen? Dan is delen - met collega's, vriendinnen, zussen, tantes of moeders - heel erg fijn!

Bij voorbaat heel veel dank voor uw tijd en hulp!

Appendix B: Scenario and Letters

Stelt u zich het volgende voor:

U bent uitgenodigd om deel te nemen aan het **gepersonaliseerde** bevolkingsonderzoek naar borstkanker. Dit bevolkingsonderzoek wordt afgestemd op uw **persoonlijke risico** op het krijgen van borstkanker.

Er zijn meerdere manieren waarop het bevolkingsonderzoek gepersonaliseerd kan worden. Uw risico zou bijvoorbeeld kunnen bepalen hoe vaak u langs mag komen voor het bevolkingsonderzoek. Op deze manier kan het bevolkingsonderzoek **op maat** worden aangeboden.

Vorige week heeft u een formulier ingevuld dat gebruikt is om uw persoonlijke borstkankerrisico te bepalen. In onderstaande brief krijgt u voor het eerst de **uitslag van deze risicobepaling**.

Lees de inhoud van deze naar u gestuurde brief rustig door en beantwoord hierna enkele vragen.

Numerical-Only – 7%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

Hieruit is gebleken dat uw persoonlijke risico op borstkanker 7% (7 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).

Numerical-Only – 28%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

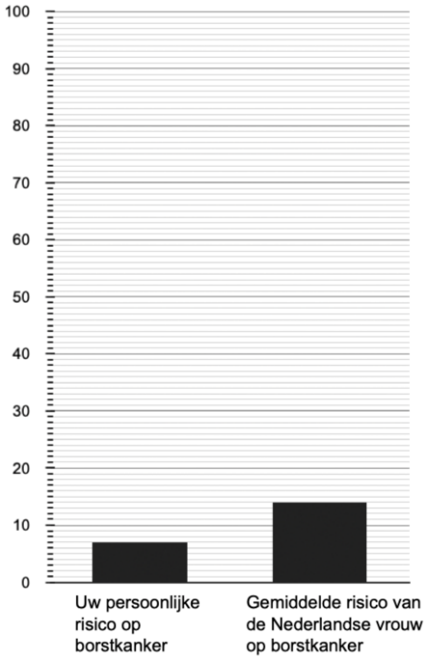
Hieruit is gebleken dat uw persoonlijke risico op borstkanker 28% (28 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).

Numerical and Bar Graph – 7%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

Hieruit is gebleken dat uw persoonlijke risico op borstkanker 7% (7 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).

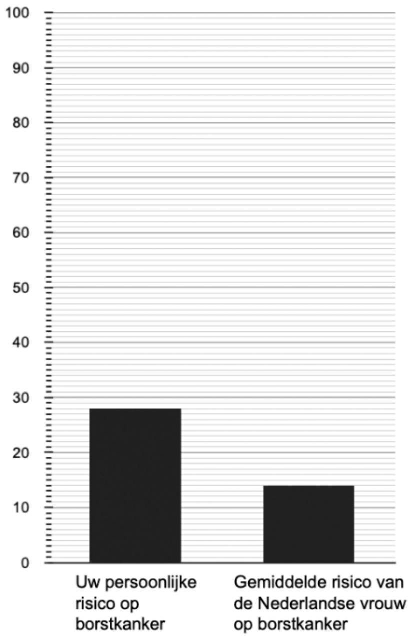


Numerical and Bar Graph – 28%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

Hieruit is gebleken dat uw persoonlijke risico op borstkanker 28% (28 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).

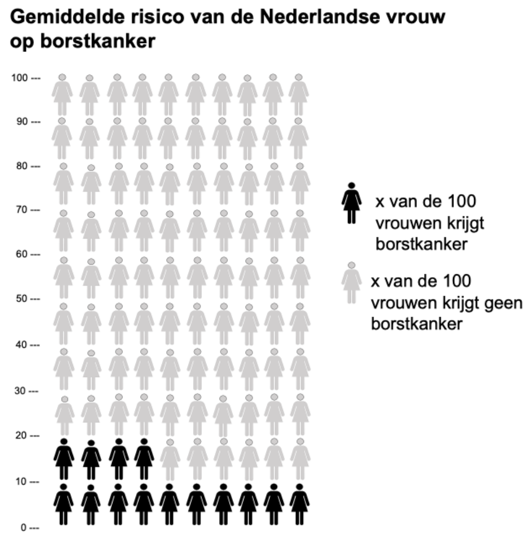


Numerical and Icon Array – 7%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

Hieruit is gebleken dat uw persoonlijke risico op borstkanker 7% (7 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).



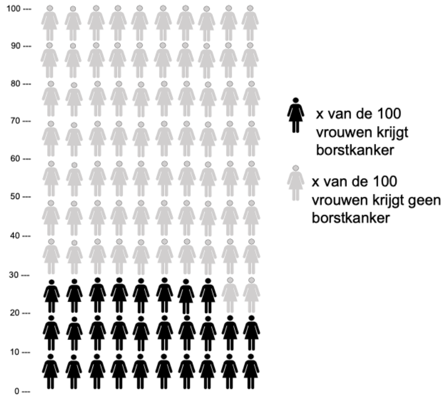
Numerical and Icon Array – 28%

Middels deze brief informeren wij u over de uitslag van uw risicobepaling.

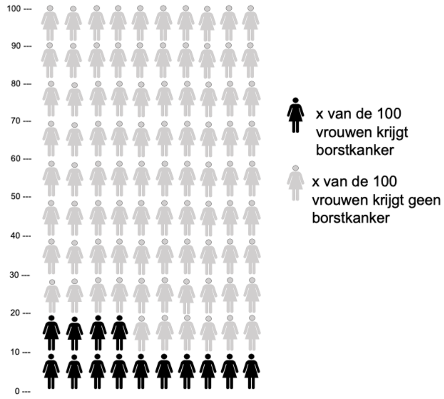
Op basis van factoren zoals uw leeftijd, geslacht, leeftijd van eerste menstruatie, leeftijd van eerste zwangerschap en familiehistorie, is uw persoonlijke risico op borstkanker gedurende uw leven geschat.

Hieruit is gebleken dat uw persoonlijke risico op borstkanker 28% (28 op 100 vrouwen) is. Het gemiddelde risico van de Nederlandse vrouw op borstkanker is 14% (14 op 100 vrouwen).

Uw persoonlijke risico op borstkanker



Gemiddelde risico van de Nederlandse vrouw op borstkanker



Appendix C: Scales and Questions

Cursive = original item.

*= only shown in the second measurement.

Accuracy of Perceived Risk (Schapira et al., 2004)

What do you think your personal risk is of getting breast cancer in your lifetime?

Please answer on a scale of 0% to 100%.

(**Op basis van de informatie in deze brief:***) Hoe hoog schat u uw persoonlijke risico of kans in op het krijgen van borstkanker gedurende uw leven, op een schaal van 0-100%?

0 = geen risico op het krijgen van borstkanker

100 = helemaal zeker dat ik borstkanker krijg

Perceived Risk

(**Op basis van de informatie in deze brief:***) Hoe groot of klein acht u uw risico op het krijgen van borstkanker gedurende uw leven?

( ) Heel klein

( ) Klein

( ) Niet klein, niet groot

( ) Groot

( ) Heel groot

(**Op basis van de informatie in deze brief:***) Hoe waarschijnlijk denkt u dat het is dat u gedurende uw leven borstkanker krijgt?

( ) Zeer onwaarschijnlijk

( ) Onwaarschijnlijk

( ) Niet onwaarschijnlijk, niet waarschijnlijk

( ) Waarschijnlijk

( ) Zeer waarschijnlijk

Breast Cancer Worry (Lipkus et al., 2005)

1. *How worried are you of getting breast cancer in your lifetime?*

   Not at all (1) (2) (3) (4) (5) (6) Extremely

   (**Op basis van de informatie in deze brief:***) Hoe bezorgd bent u over het krijgen van borstkanker gedurende uw leven?

   Helemaal niet (1) (2) (3) (4) (5) (6) Heel erg

2. *How fearful are you of getting breast cancer in your lifetime?*

   Not at all (1) (2) (3) (4) (5) (6) Extremely

   (**Op basis van de informatie in deze brief:***) Hoe angstig bent u voor het krijgen van borstkanker gedurende uw leven?

   Helemaal niet (1) (2) (3) (4) (5) (6) Heel erg

3. *How anxious are you of getting breast cancer in your lifetime?*

   Not at all (1) (2) (3) (4) (5) (6) Extremely

   (**Op basis van de informatie in deze brief:***) Hoe ongerust bent u over het krijgen van borstkanker gedurende uw leven?

   Helemaal niet (1) (2) (3) (4) (5) (6) Heel erg

Dutch Translation of Subjective Numeracy Scale (Vromans et al., 2020)

Geef per vraag aan **hoe goed u bent in het doen van de volgende dingen**:

1. Hoe goed bent u in het rekenen met breuken?

   Helemaal niet goed (1) (2) (3) (4) (5) (6) Heel erg goed

2. Hoe goed bent u in het rekenen met percentages?

   Helemaal niet goed (1) (2) (3) (4) (5) (6) Heel erg goed

3. Hoe goed kunt u een fooi van 15% berekenen?

   Helemaal niet goed (1) (2) (3) (4) (5) (6) Heel erg goed

4. Hoe goed kunt u de prijs van een t-shirt berekenen als er 25% korting afgaat?

   Helemaal niet goed (1) (2) (3) (4) (5) (6) Heel erg goed

Geef per vraag aan wat bij u van toepassing is.

5. Als u de krant leest, hoeveel hebt u dan aan tabellen en grafieken die bij een artikel horen?

   Helemaal niets (1) (2) (3) (4) (5) (6) heel erg veel

6. Stel: mensen vertellen u over de kans dat iets zal gebeuren. Wilt u dan dat ze **woorden** gebruiken (bijvoorbeeld “het gebeurt zelden”) of **getallen** (bijvoorbeeld “de kans is 1%)?

   Ik geef altijd de voorkeur aan woorden (1) (2) (3) (4) (5) (6) Ik geef altijd de voorkeur aan getallen.

7. Stel: u luister naar het weerbericht. Hoort u dan liever voorspellingen in **procenten** (bijvoorbeeld “vandaag is de kans op regen 20%) of voorspellingen in **woorden** (bijvoorbeeld “vandaag is er een kleine kans op regen”)?

   Ik geef altijd de voorkeur aan percentages (1) (2) (3) (4) (5) (6) Ik geef altijd de voorkeur aan woorden.

8. Hoe **vaak** heeft u iets aan informatie in de vorm van getallen?

   Nooit (1) (2) (3) (4) (5) (6) Heel vaak

Appendix D: Assumption Evaluations and Results of Analyses

Age

The mean age for three of the six groups were not normally distributed (z-score $skewness_{numerical\ 7\%}$ = 2.05, z-score $skewness_{bar\ graph\ 7\%}$ = 2.84, z-score $skewness_{icon\ array\ 7\%}$ = 2.49). For this reason, bootstrapping was performed. The assumption of homogeneity of variances was met. Levene's test was not significant: $F(5, 397) = 1.78$, $p = .117$). The ANOVA showed no significant effect: $F(5, 397) = 1.16$, $p = .328$.

Subjective Numeracy

The subjective numeracy scores for two of the six groups were not normally distributed (z-score $skewness_{numerical\text{-}only\ 7\%}$ = -2.54, z-score $kurtosis_{numerical\text{-}only\ 28\%}$ = -2.03). For this reason, bootstrapping was performed. The assumption of homogeneity of variances was met. Levene's test was not significant: $F(5, 397) = 0.92$, $p = .466$). The ANOVA showed no significant effect: $F(5, 397) = 0.39$, $p = .856$.

Education Level

All assumptions were met. There was no significant association between education level and the six conditions (χ^2 (10) = 10.48, $p = .400$). None of the cells contributed significantly to the overall test statistic.

Relatives With Breast Cancer

All assumptions were met. There was no significant association between education level and the six conditions (χ^2 (15) = 7.28, $p = .949$). None of the cells contributed significantly to the overall test statistic.

First Manipulation Check

All assumptions were met. There was no significant association between scores on the first manipulation check and the six conditions (χ^2 (5) = 5.35, $p = .375$). None of the cells contributed significantly to the overall test statistic.

Second Manipulation Check

The baseline second manipulation check scores for one of the six groups were not normally distributed (z-score $skewness_{bar\ graph\ 7\%} = 2.36$). For this reason, bootstrapping was performed. The assumption of homogeneity of variances was met. Levene's test was not significant: $F(5, 397) = 1.34$, $p = .246$).

The ANOVA showed a significant effect: $F(5, 397) = 7.09$, $p < .001$, $\eta^2 = .08$. A Post Hoc Tukey-HSD analysis revealed that multiple groups differed from one another. The *p*-values, bootstrapped confidence intervals and mean differences are presented in Table 7.

Table D1

Results of Post Hoc Tukey-HSD Test

	Numerical-only 7%	Bar graph 7%	Icon array 7%
Numerical-only 28%		*p* = .015 BCa 95% CI [0.20, 0.79] *Mdif* = 0.50	
Bar graph 28%	*p* = .010 BCa 95% CI [0.20, 0.78] *Mdif* = 0.50	*p* = .001 BCa 95% CI [0.31, 0.94] *Mdif* = 0.64	*p* = .005 BCa 95% CI [0.24, 0.85] *Mdif* = 0.54
Icon array 28%	*p* = .016 BCa 95% CI [0.20, 0.74] *Mdif* = 0.47	*p* = .001 BCa 95% CI [0.30, 0.91] *Mdif* = 0.61	*p* = .008 BCa 95% CI [0.23, 0.81] *Mdif* = 0.51

Baseline Perceived Risk

The baseline risk perception scores for all six groups were not normally distributed (z-score $skewness_{numerical\text{-}only\ 7\%} = 3.78$, z-score $kurtosis_{numerical\text{-}only\ 7\%} = 2.01$, z-score $skewness_{numerical\text{-}only\ 28\%} = 2.86$, z-score $skewness_{bar\ graph\ 7\%} = 2.92$, z-score $kurtosis_{bar\ graph\ 28\%} = 2.42$, z-score $skewness_{icon\ array\ 7\%} = 2.65$, z-score $skewness_{icon\ array\ 28\%} = 2.32$, z-score kurtosis$_{\text{icon array 28\%}}$ = 2.41). For this reason, bootstrapping was performed. The assumption of homogeneity of variances was met. Levene's test was not significant: $F(5, 397) = 0.94$, p = .454). The ANOVA showed no significant effect: $F(5, 397) = 0.83$, $p = .531$.

Baseline Breast Cancer Worry

The baseline breast cancer worry scores for two of the six groups were not normally distributed (z-score skewness$_{\text{bar graph 28\%}}$ = 2.80, z-score skewness$_{\text{icon array 7\%}}$ = 2.16). For this reason, bootstrapping was performed. The assumption of homogeneity of variances was met. Levene's test was not significant: $F(5, 397) = 1.04$, $p = .394$). The ANOVA showed no significant effect: $F(5, 397) = 1.53$, $p = .180$.

Assumption Evaluations Two-Way MANCOVA

The Kolmogorov-Smirnov test of normality was significant for all dependent variables and the covariate, meaning that the data were not normally distributed (accuracy of perceived risk: $D(403) = 0.27$, $p < .001$, breast cancer worry: $D(403) = .20$, $p < .001$, perceived risk: $D(403) = 0.21$, $p < .001$, numeracy: $D(403) = 0.06$, $p < .001$). The Box's M of 41.50 indicates that the homogeneity of covariance matrices across groups is assumed ($F(30, 348754.37) = 1.36$, $p = .092$), since the p-value is higher than .001. However, the assumption of homogeneity of variance was not met for 'breast cancer worry' ($F(5, 397) = 2.61$, $p = .024$). The assumption of linearity was generally met, although some combinations of variables did not show elliptical shapes. The assumption of multicollinearity was met: None of the correlations between the dependent variables was above $r = .90$.

Assumption Evaluations Three-Way MANOVA (Exploratory Analysis)

The Kolmogorov-Smirnov test of normality was significant for all dependent variables, meaning that the data were not normally distributed (accuracy of perceived risk$_{\text{less numerate}}$: $D(207) = 0.21$, $p < .001$, breast cancer worry$_{\text{less numerate}}$: $D(207) = 0.20$, $p < .001$, perceived risk$_{\text{less numerate}}$: $D(207) = 0.24$, $p < .001$, accuracy of perceived risk$_{\text{highly}}$

$_{\text{numerate}}$: $D(196) = 0.34$, $p < .001$, breast cancer worry$_{\text{highly numerate}}$: $D(196) = 0.19$, $p < .001$, perceived risk$_{\text{highly numeracy}}$: $D(196) = 0.19$, $p < .001$).

Less Numerate

The Box's *M* of 41.10 indicates that the homogeneity of covariance matrices across groups is assumed ($F(30, 86410.27) = 1.32$, $p = .115$), since the *p*-value is higher than .001. The assumption of homogeneity of variance was met for all dependent variables. The assumption of linearity was generally met, although some combinations of variables did not show elliptical shapes. The assumption of multicollinearity was met: None of the correlations between the dependent variables was above $r = .90$.

Highly Numerate

The Box's *M* of 43.67 indicates that the homogeneity of covariance matrices across groups is assumed ($F(30, 68216.57) = 1.40$, $p = .074$), since the *p*-value is higher than .001. However, the assumption of homogeneity of variance was not met for breast cancer worry ($F(5, 190) = 2.93$, $p = .014$). The assumption of linearity was generally met, although some combinations of variables did not show elliptical shapes. The assumption of multicollinearity was met: None of the correlations between the dependent variables was above $r = .90$.

Less Versus Highly Numerate

The Box's *M* of 109.05 indicates that the homogeneity of covariance matrices across groups can be assumed ($F(66, 146255.63) = 1.59$, $p = .002$), although the *p*-value is only slightly higher than .001. However, the assumption of homogeneity of variance was not met for breast cancer worry ($F(11, 391) = 2.12$, $p = .018$) and for accuracy of perceived risk ($F(11, 391) = 1.90$, $p = .038$).

Appendix E: Table and Figures Exploratory Analysis

Table E1

Means and Standard Deviations for (Accuracy of) Perceived Risk and Breast Cancer Worry per Condition, Divided by Numeracy

Less numerate (score 0 - 4.25)	Numerical-only		Bar graph		Icon array	
	7%	28%	7%	28%	7%	28%
	n = 32	*n* = 37	*n* = 35	*n* = 39	*n* = 30	*n* = 34
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Perceived risk (%)	23.10 (22.04)	28.78 (17.37)	19.43 (16.74)	34.77 (18.45)	19.83 (16.80)	32.53 (18.47)
Accuracy of perceived risk[a]	16.09 (22.04)	0.78 (17.37)	12.43 (6.77)	6.77 (18.45)	12.83 (16.80)	4.53 (18.47)
Perceived risk[b]	2.56 (0.70)	2.85 (0.77)	2.37 (0.70)	2.92 (0.65)	2.42 (0.53)	2.87 (0.77)
Breast cancer worry[c]	2.45 (1.10)	2.91 (1.13)	2.36 (1.01)	2.91 (1.06)	2.19 (0.87)	3.11 (1.13)
Highly numerate (score 4.26 - 6)	**Numerical-only**		**Bar graph**		**Icon array**	
	7%	28%	7%	28%	7%	28%
	n = 41	*n* = 29	*n* = 27	*n* = 27	*n* = 35	*n* = 37
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Perceived risk (%)	16.17 (16.05)	31.38 (14.29)	14.59 (14.51)	33.19 (15.88)	12.34 (14.03)	33.35 (14.34)
Accuracy of perceived risk[a]	9.17 (16.05)	3.38 (14.29)	7.59 (14.51)	5.19 (15.88)	5.34 (14.03)	5.35 (14.34)
Perceived risk[b]	2.11 (0.73)	2.97 (0.52)	2.07 (0.60)	2.93 (0.65)	2.17 (0.67)	3.05 (0.61)
Breast cancer worry[c]	2.17 (0.88)	3.30 (1.04)	2.00 (0.65)	2.99 (1.25)	1.98 (0.88)	3.43 (0.97)

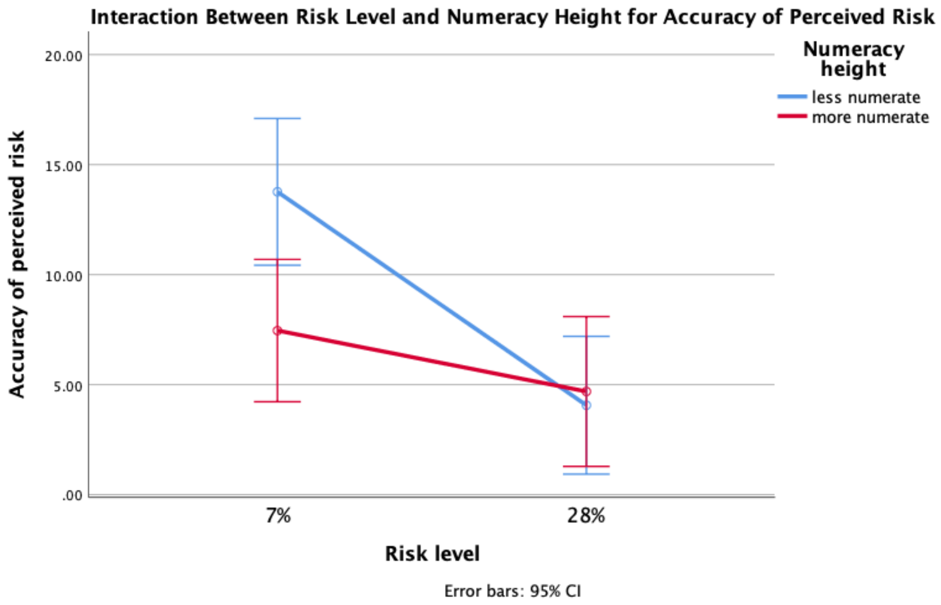
[a] Scores closer to zero are more accurate.

[b] Measured on a 5-point scale.

[c] Measured on a 6-point scale

Figure E1

Interaction Effect for Accuracy of Perceived Risk



Note. Scores closer to zero are more accurate.

Figure E2

Interaction Effect for Perceived Risk

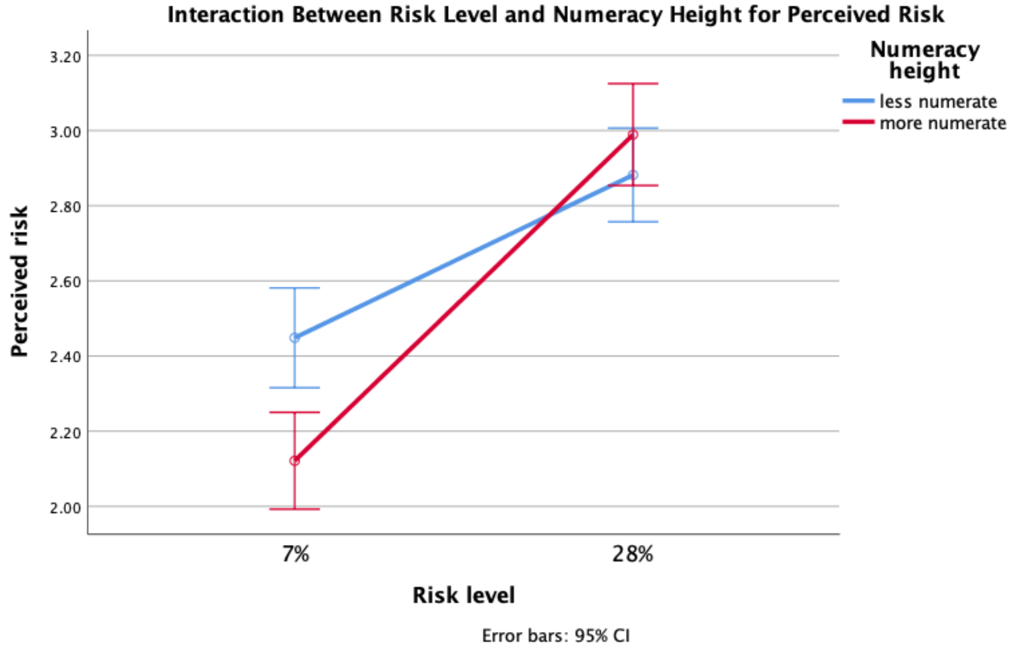


Figure E3

Interaction Effect for Breast Cancer Worry