Published in final edited form as:
Risk Anal. 2009 November ; 29(11): 1549–1565. doi:10.1111/j.1539-6924.2009.01299.x.

The Effect of Graphics on Environmental Health Risk Beliefs, Emotions, Behavioral Intentions and Recall

Dolores J. Severtson* and **Jeffrey B. Henriques**
UW-Madison School of Nursing

Abstract

Lay people have difficulty understanding the meaning of environmental health risk information. Visual images can use features that leverage visual perception capabilities and semiotic conventions to promote meaningful comprehension. Such evidence-based features were employed to develop two images of a color-coded visual scale to convey drinking water test results. The effect of these images and a typical alphanumeric (AN) lab report were explored in a repeated measures randomized trial among 261 undergraduates. Outcome measures included risk beliefs, emotions, personal safety threshold, mitigation intentions, the durability of beliefs and intentions over time, and test result recall. The plain image conveyed the strongest risk message overall, likely due to increased visual salience. The more detailed graded image conveyed a stronger message than the AN format only for females. Images only prompted meaningful risk reduction intentions among participants with optimistically biased safety threshold beliefs. Fuzzy trace theory supported some findings as follow. Images appeared to promote the consolidation of beliefs over time from an initial meaning of safety to an integrated meaning of safety and health risk; emotion potentially shaped this process. Although the AN report fostered more accurate recall, images were related to more appropriate beliefs and intentions at both time points. Findings hinted at the potential for images to prompt appropriate beliefs independent of accurate factual knowledge. Overall, results indicate that images facilitated meaningful comprehension of environmental health risk information and suggest foci for further research.

Keywords

risk communication; graphics; fuzzy trace theory; cognitive bias; visual information; decision making

1. INTRODUCTION

Our interest in images was prompted by a previous study of private well owners' responses to well water test results for arsenic.(1) Findings showed moderately strong positive relationships between water test results, safety beliefs, and mitigation behavior. The authors concluded that a test result compared to a safety standard provides concrete evidence of an unseen risk and is therefore more powerful than abstract risk information. Despite this, more

*Address correspondence to Box 2455 H6/236 CSC, 600 Highland Ave., Madison WI, 53792; tel: 608-263-5311; fax: 608-263-5332; djsevert@wisc.edu .

SUPPORTING INFORMATION The following supporting information is available via a link in the online table of contents for this journal.
Fig. S1. Plain Red Image
Fig. S2. Graded Red Image
Table S1. Variables at Time 1 and Time 2 with Means and Standard Deviations or Frequencies by Format

than half of participants had inappropriate safety beliefs and behavior prompting a recommendation to explore how the *meaning* of water test results could be conveyed more concretely.(1) Graphical representations can make abstract information more concrete(2,3) and have been recommended for conveying environmental monitoring information as it relates to safety standards or benchmarks,(4) thus appropriate for addressing this objective.

We begin by reviewing factors that may have contributed to inappropriate safety beliefs and protective mitigation behavior. These factors shaped the variables included in the study.

1.1. Potential Barriers to Appropriate Beliefs and Behavior

The technical nature of laboratory test results and the terminology used to describe water test results may be a barrier to comprehension.(4) Half of participants in a survey study of arsenic well water testing(1) rated their water test result as less than easy to understand, and about a third of written suggestions for laboratories were requests for user friendly test results (unpublished results). Johnson(4) examined how people understood phrases used to describe water test results as they relate to drinking water standards. Typically used phrases such as "exceeds the standard" or "above the standard" were sometimes misunderstood as meaning better than the standard. Safety standards can also promote a *threshold effect* (beliefs that results just below the standard are safe while those just above are unsafe), rather than a more accurate meaning of a dose-risk continuum.(4) Risk ladder graphics(5) appear to mitigate these threshold effects.

Defensive denial can minimize threatening information. Threat seriousness and ambiguity trigger defensive beliefs that discount the seriousness or validity of a health threat, diagnosis or test result.(6) Some private well owners recalled a lower test result than their actual result and/or selected a personal safety threshold that was less stringent than the drinking water standard;(1) both would defend judgments of safe water when test results indicate otherwise. Information that conveys the meaning of test results may be less vulnerable to denial. Across two field studies of responses to cholesterol test results, participants had accurate recall of a concrete interpretation of their test results (low, borderline, high), compared to optimistic recall of their actual numeric results,(7,8) suggesting that a meaningful interpretation of test results may mitigate denial.

1.2. Visual Images as Information

Visualization can make information easier to understand. Research evidence on how visual features are perceived and understood is found across disciplines.[1] The use of features that are easily and consistently comprehended can narrow the gap between the intended and the imparted meaning of a graphic. Visual perception and cognition research sheds some light on *how* and *why* images or image features facilitate comprehension. Limited short term memory can create a bottleneck for processing information. On average, people can hold about four "chunks" of information in short term memory at a time.(9) Some "elementary" visual features, (for example position on a common scale and direction) are rapidly and accurately detected due to the innate abilities of human visual perception.(10) Elementary features, readily apprehended in parallel with other information and without the need for active information processing, do not need to be stored in short term memory for cognitive processing. This frees up cognitive capacity for processing other information features.(10)

[1]For example: visuospatial perception and cognition,(52) behavioral and communication science pertaining to environmental(2) or health risk,(3,29,53) industrial ergonomics,(45) and cartography.(54)

The strategic use of evidence-based[2] visual features may address literacy and numeracy barriers for risk communication by facilitating seemingly automatic comprehension. We illustrate this by summarizing multiple forms of evidence that support the effectiveness of a vertical scale for communicating risk. First, Cleveland and McGill(11) found that position on a common scale was the most accurately understood among ten proposed elementary perceptual features for conveying quantitative information.[3] Furthermore, Golledge(12) proposed that spatial information facilitates *configurational knowledge* about how objects are spatially related to each other, for example the direction and magnitude of difference between two points on a scale. Configurational knowledge is considered as a higher order of understanding than declarative knowledge.(12) Weinstein and others(5,13) found that *position on a vertical scale* influenced risk beliefs more than the numerical units on the scale; the spatial component of the scale was more influential than the numerical component. Configurational knowledge derived from seeing risk magnitudes on a common scale suggests why vertical scales are effective and illustrates how the appropriate application of visual features can improve comprehension.

Shading and color saturation were listed as the least accurately understood of Cleveland and McGill's ten elementary visual features.(10) Cartographic convention is that sequentially darker shades of a color convey increasing magnitude on maps.(14) Color gradients may have utility for conveying quantity for other graphics.

Semiotics is the study of signs. MacEachren posits that semiotic features imbue images with meaning.(15) Signs (icons, symbols) decrease load on limited short term memory because they rely on meaningful visual conventions that are readily available from long term memory. These conventions may map to the physical world (blue to depict water bodies on a map), or to socio-cultural conventions.(15) Widely used symbols for communicating risk include "stop sign" colors of green for safe, yellow for caution, and red for danger.(16) Signs are more effective when they depict concrete rather than abstract concepts(17) and communicate meaning without the need for a legend.(18) Signs can increase comprehension and recall; warning messages in an instructional manual that included a pictorial icon were better comprehended and more accurately recalled than those that did not.(19)

1.3. Theory

Visual perception *gestalt* pertains to the idea that humans perceive images as integrated wholes rather than composites of various features such as lines and curves.(20) An image that consists of a variety of evidence-based features is likely to convey a consolidated risk message rather than multiple messages derived from the specific features. This consolidated message may result in more global gist-based than detailed fact-based understandings. Reyna and Brainerd(21) propose that people have a preference for fuzzy rather than precise information processing, with a goal of deriving meaning (gist) from information rather than precise details. Gist understandings are held to be more cognitively sophisticated and applicable to decision-making than detailed fact-based knowledge. These authors propose that cognition results in parallel but separate memory traces for verbatim and gist-based knowledge. These separate traces explain disparities between recalled facts and gist derived from the same experience. Verbatim facts tend to fade over time, while gist becomes consolidated and integrated with overarching principles.(22)

[2]The concept of *evidenced-based* is borrowed from the term *evidence-based practice*. It conveys the idea that practice should be based on the most current, valid, and reliable research evidence that is available.(55)

[3]These proposed features were ordered from most to least accurate as: (1) position along a common scale, (2) positions along nonaligned scales, (3) length, direction, angle, (4) area, (5) volume, curvature, (6) shading, color saturation.(11)

Images are representations. Zhang and Norman(23) called for understanding theoretical relationships between external graphical representations and internal mental representations to explicate how graphical representations are used for learning and making decisions. Lipkus(3) observed that decisions and behavior may be shaped more by the *meaning* derived from graphics than specific facts. People's risk beliefs embody the concept of meaning. In the area of health behavior, perceived risk is often conceptualized as a function of risk likelihood and severity beliefs. Problem seriousness, a more global appraisal of risk, is a function of beliefs about the likelihood and severity of a health risk and a strong predictor of mitigation behavior.(24) Fuzzy trace theory may have utility for understanding responses to image information.(3,25) Measuring how images shape beliefs (akin to gist) that are known to be related to behavior may provide insights into the potentially actionable meaning that people derive from images. Images are more likely to engender emotions than equivalent alphanumeric information.(2) For example, even non-contextual color is related to emotion.(26) Emotion appears to have a role in the development of gist(27) and is tightly integrated with cognition(28) thus warning colors would be expected to induce both stronger beliefs and emotions.

Graph features that promote accurate comprehension do not necessarily promote appropriate behavior, and vice versa.(29) Therefore, assessing how various graphical formats impact recalled facts along with beliefs and feelings that are associated with protective health behavior may provide theoretical insights into *how* images influence behavior. For example, Chua and others(30) found that participants assigned to a graphic (compared to a numeric) display of product safety information were more likely to recommend the safer product to a friend. Emotion and risk likelihood beliefs had stronger effects on outcomes among participants assigned to the graphic rather than numeric display suggesting potential mechanisms for the image effect. Participants assigned to numerical information had more accurate factual recall than those assigned to graphical information suggesting that beliefs play a greater role than accuracy in shaping actionable behavior.(30)

1.4. Development of Water Test Result Images and Study Aims

Evidence-based visual features were applied to create two versions of a color coded scale to convey water test results, the safety meaning of results, and a brief message that conveyed agency-based recommendations for responding to test results at the time of the study. The plain image included the following features:

- A common vertical scale conveyed the dose-risk magnitude(5,11,13) and spatial configuration of how the water test result was related to the safety standard.(12)
- Conventional warning colors of green (safe), yellow (caution), red (danger), and black (fatal)(31) were used to convey the safety implications of test results.
- Concrete action-oriented information is more likely to prompt protective health behavior change than abstract information.(32) Short action messages of “No recommended action”, Test well yearly”, “Stop drinking”, and “Stop showering” were placed in the green, yellow, red and black blocks respectively.
- A pictorial icon(19) depicted a glass of water, the universal symbol for drinking water, overlaid with the universal symbol for stop. It was inserted into the red *stop drinking* block of the scale.

The graded color scale used six incrementally darker red gradients(14) to reinforce the dose-risk message for the *stop drinking* section of the scale, in addition to the above features. An alphanumeric (AN) laboratory test result that included the drinking water standard was constructed that typified the report provided by laboratories in the state where this study was conducted. Images provided recommended actions but the AN report did not. See footnote [4]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

for details. The three versions are illustrated in Figures 1 – 3. Color versions are available via a link in the online journal table of contents.

Our primary research goal was to explore the influence of these three formats (Figures 1-3) of the well test result for two hypothetical risk doses (well test results) on study outcomes. Specific research questions were:

1. What is the influence of format on the personal safety threshold?
2. What is the influence of format on the intentions to mitigate risk and relative risk reduction?
3. How does format influence beliefs and intentions for the model in Figure 4?
4. What are the within-subject influences of format on *specific* and *global risk beliefs*, *emotions*, and *risk mitigation intentions* over time?
5. How does format influence the accuracy of dose-risk beliefs? Does the graded image strengthen the dose-risk message compared to the plain image?
6. How does format influence recall accuracy of the dose and dose-standard relationship?
7. How does format influence the appropriateness of safety beliefs and risk mitigation intentions among participants with incorrect dose and dose-standard relationship recall?

These research aims address several of the visual risk communication research recommendations offered by Lipkus:(3) using randomized trials to test formats against each other, examining mediating mechanisms for how formats shape outcomes of interest, and building theoretical understandings for how visual features work.

2. METHODS

2.1. Sample and Study Design

Participants were 350 students in an undergraduate psychology class at a large mid-western university. This survey study used a 3 × 2 repeated measures design for the three different formats (AN, plain image, graded image) and two different water contaminant doses (12 and 40 ppb) that were over a drinking water standard of 10 ppb. These doses were selected to explore: (1) the influence of the plain compared to the graded color image (lower and higher doses would fall into lighter and darker red gradations), and (2) the effect of format relative to, and potentially interacting with, dose.

2.2. Surveys and Procedure

Survey 1 consisted of 12 items to measure *specific* (likelihood and severity) and *global* (safety and problem seriousness) *risk beliefs*, *numerical dose-risk estimates*, *emotions* (concern and distress), behavioral *risk mitigation intentions* (drink less untreated water), and defensive denial (*personal safety threshold*). A measure of *skepticism* about contaminant-related health effects was included for exploratory purposes since some members of the public are generally skeptical about health-related advice.(33) Survey 2 consisted of the same items in survey 1 but without the *personal safety threshold* and *dose/risk belief*

[4]The images included risk mitigation information (“stop drinking” message and icon) not included with the AN format. The AN format included *LOQ*, *analysis method*, *date* and *analyst* information, but the images did not. The interpretive message on the AN image (the test result was over the drinking water standard of 10 ppb) was equivalent to the configurational message conveyed on the image by the position of the well test result relative to the drinking water standard.

measures. Additional items measured *dose recall* (< 10, 10 – 50, > 50), and *recall* of whether the dose was *under or over the standard*.

Participants were randomly assigned to 1 of the 3 formats by uniformly mixing the study packets prior to distribution among students who volunteered to participate in the study at the end of class. A scenario and fact sheet were provided to give participants a standard context for interpreting the fictitious well test result. (Scenario: They were provided with well water test results for their family residence of 30 years. A fictitious "fact" sheet provided common health effects and statements that introduced ambiguity for the drinking water standard. Details are in footnote [5].) A week later, students received only survey 2 (without test result) and were asked to answer based on their current beliefs about the lab test result they had received the previous week.[6]

2.3. Analysis

SPSS (Version 15.0) was used for all reported analyses unless otherwise noted. Alpha values < 0.10 are reported. Researchers suggest relaxing alpha for exploratory intervention studies to guard against rejecting findings that may provide useful insights.(34) Analysis methods are in the results section to illustrate how findings from research questions 1-2 (showing gender and threshold bias as potential moderators) led to exploring format effects across moderator defined subgroups for remaining questions. Moderators explain the circumstances under which interventions are more or less effective.(35) Gender may moderate format effects because females are more likely to attend to and recall details including visual details,(36) are more risk averse,(37) and perform somewhat worse on some math and spatial cognition tests(38) than males.

3. RESULTS

Both surveys were completed by 261 (74.6%) students. The mean age was 19.5*(1.2)* years and 62% were female. Means and standard deviations for variables at time time 1 and 2 across formats are available in an online table at Table S1 via a link in the online journal table of contents. Due to the multiple analyses across subgroups, null results are not reported but available from the corresponding author. Results are organized by research question (RQ).

3.1. RQ 1

What is the influence of format on the personal safety threshold?

Format had no effect on the *personal safety threshold*; $F(2, 259) = 0.47$, $p = .625$. Being male or having a higher dose were related to selecting a higher *personal safety threshold*; gender $F(1, 260) = 9.31$, $p = .003$; dose $F(1, 260) = 8.41$, $p = .004$. There were no interaction effects.

3.1.1. Analysis Decision and Rationale—We created a dichotomized version of the *personal safety threshold* called *optimistic threshold bias* (OTB), and specified OTB as an independent variable in subsequent analyses. OTB denoted personal safety threshold beliefs

[5]Participants were asked to imagine they had graduated and were returning home to take over the family business. The private well that supplied drinking water to their home of 30 years had been recently tested for manganese and the results were provided to them. The brief list of facts, based on arsenic but specified for *manganese*, provided common health effects and information that introduced ambiguity that was similar to the ambiguity that existed for arsenic at the time of the study - that the drinking water standard was recently revised from 50 ppb to 10 ppb and that while most experts agreed that the standard should be lower, a few did not.

[6]Students were debriefed regarding the fictitious nature of the information via a printed information sheet that was provided when they turned in the survey. The study protocol was reviewed by the university's IRB and approved as minimal risk.

that were aligned with or more protective than the drinking water standard (≤ 10 = no OTB), or less protective (> 10 = OTB) and thus supporting optimistically biased safety beliefs. This decision was supported by: (1) null effects of format on the personal safety threshold, (2) a considerable body of work showing that optimistically biased beliefs about health threats are prevalent and often quite stable(39) and embody a variety of cognitions related to optimism,(40) and (3) a strong relationship between OTB and the dichotomized *skepticism* variable [$\chi2(1, n=261) = 18.783$, $p = <.001$], but no relationship between OTB and dose [$\chi^2(1, n=261) = 0.89$, $p = .347$]. ANOVA showed that skepticism was not influenced by dose or format[7] suggesting it was not defensively induced by threat.

3.2. RQ 2

What is the influence of format on the intentions to mitigate risk relative risk reduction?

A full factorial 4-way ANOVA with Sidak corrections for multiple *post hoc* comparisons was used to examine the influence of format on risk mitigation intentions in the presence of dose, gender, and OTB.

Format had significant effects on *risk mitigation intentions* at both time points (Table I). The plain image was related to stronger *risk mitigation intentions* than the AN format at both points in time (Table II). Participants with no OTB, those assigned to a stronger dose, and those who were female had stronger *risk mitigation intentions*. The OTB by gender interaction occurred because among participants with OTB, females had stronger *risk mitigation intentions* compared to males at both time points [time 1: $t(160) = 3.06$, $p = .003$; time 2: [$t(160) = 2.82$, $p = .005$] but among those with no OTB, these gender effects were reversed although not significant [time 1: [$t(97) = -1.36$, $p = .176$; time 2 $t(97) = -1.32$, $p = .191$]. The time 1 format by gender interaction was explored via separate ANOVAs for males and females and Sidak-corrected post hoc format comparisons (Table 2). The plain compared to AN format promoted stronger *risk mitigation intentions* across both genders (time 1 and 2), but the graded image only had this effect among females. Among males, intentions were stronger for the plain compared to graded image, but not among females. These striking and persistent gender effects for the graded image and interaction effects with OTB led to conducting further analysis across subgroups defined by format, gender and OTB.

We examined how the formats were related to a meaningful difference in risk reduction between formats by calculating relative risk reduction (RRR); RRR is the relative event rate for a treatment group compared to the rate for a comparison group.(41) First, intention to stop drinking one's water (our measure of *risk mitigation intention*) was dichotomized (*drink as usual/a little less* and *drink less/very little/stop drinking*). Then, rates of RRR between format pairs split by gender and OTB were compared using the online Clinical Significance Calculator.(42) We defined meaningful risk reduction as RRR > 25%.

Results are in Table III. Format only produced RRR > 25% among participants with OTB; this occurred at time 1 and 2 for either image among females but only the plain format among males.

3.3. RQ 3

How does format influence beliefs and intentions for the proposed model of causal influence in Figure 4?

---

[7]Independent variables on *skepticism*: format $F(2, 259) = 0.69$, $p = .503$, dose $F(1, 260) = 0.93$, $p = .336$, gender $F(1, 260) = 2.30$, $p = .130$.

We conducted three path analyses (using M-plus(43)) to examine the effect of each image relative to the AN format and the plain relative to the graded format for the model depicted in Figure 4. Interactions among variables at the front of the model were only included if consistently significant across time. The indirect effects through each *belief* variable or *distress* were summed for that variable. For a path through two mediating variables (for example *severity* and *safety*), the indirect path effect was added to the sum for each of the two variables. Thus, summed indirect effects were greater than the total indirect effect for each variable. Table IV shows model R^2, fit indices, and y-standardized direct and indirect effects(44),[8]

The plain format (compared to AN) had substantial positive effects on *risk mitigation intentions* at both time points. Total effects were roughly split between direct and indirect effects. At time 1, 70% of indirect effects were via stronger *safety beliefs*, but only 28% at time 2. Time 2 indirect effects were distributed among *specific* and *global beliefs*. The plain image was related to stronger *global beliefs* (*safety* and *problem seriousness*) and *distress* At time 1 these were nearly all direct effects, but at time 2 about half were indirect effects via *severity* and *likelihood*.

The graded format (compared to AN and to plain) had substantial gender by format effects but no main effects. Among females (null for males), the graded image (compared to AN) was related to stronger *risk mitigation intentions* at both time points; predominantly as indirect effects (84% at time 1 and 65% at time 2). At time 1, 67% of indirect effects were via stronger *safety beliefs*, but only 25% at time 2. Time 2 indirect effects were evenly distributed across *specific* and *global beliefs* and *distress*. The graded image was related to stronger *global beliefs* and *distress* at both time points, predominantly as direct effects, especially at time 1.

Among males (null for females), the plain format (compared to graded) had moderate positive effects on *risk mitigation intentions*, predominantly as direct effects (75% at time 1 and 80% at time 2). There were no format effects on beliefs for this contrast.

3.4. RQ 4

What are the within-subject influences of format on *specific* and *global risk beliefs*, *emotions*, and *risk mitigation intentions* over time?

Within-subject change (Δ) over time was computed by subtracting time 2 from time 1 measures. ANOVA explored format effects on Δ time measures for *specific* and *global risk belief* variables, *emotions*, and *risk mitigation intentions* in the presence of dose, gender, format, and OTB.

There were only main format effects for Δ *severity* [$F(2, 259) = 3.31$, $p = .038$], and Δ *distress* [$F(2, 259) = 2.59$, $p = .077$]. In addition, there were multiple format interaction effects on Δ *distress*: format by dose [$F(2, 259) = 3.36$, $p = .036$]; format by gender [$F(2, 259) = 3.06$, $p = .049$]; format by OTB [$F(2, 259) = 2.68$, $p = .070$], and format by dose by gender by OTB [$F(2, 259) = 4.15$, $p = .017$]. Additional ANOVA's split by dose, gender and OTB showed format effects only among low dose, male, or participants with no OTB.

Significant format effects were further explored using Chi-square between format pairs for 3 Δ time subgroups defined by Δ severity beliefs (weaker, no change, stronger), see Table V.

---

[8]Y-standardized values should be used for binary covariates because a standard deviation change of a binary variable is not meaningful. The y-standardized coefficient is interpreted as the change in y-standard deviation units when x changes from zero to one.(44)

Severity became stronger over time for a larger proportion of participants assigned to either image compared to the AN format.

To further understand these effects, mean *severity* and *distress* values at time 1 and time 2 were compared across each of the Δ time subgroups. Exploratory analysis across format, OTB and gender subgroups found consistent patterns among all but males assigned to the graded format. Table VI shows results for the whole sample excluding males with the graded format (n=227). Strong and weak mean *severity* at time 1 reversed to become respectively weak and strong at time 2. Mean d*istress* was roughly equivalent across Δ subgroups and remained roughly the same from time 1 to time 2. For the stronger and weaker Δ severity subgroups, time 1 mean *distress* was substantially stronger and weaker than time 1 mean *severity*.

3.5. RQ 5

How does format influence dose-risk beliefs? Does the graded image strengthen the dose-risk message compared to the plain image?

The numerical nature of dose-risk beliefs was assessed using a categorical *dose-risk estimate* item dichotomized into more accurate and less accurate answers.[9] A 3 × 2 Chi-square analysis across males and females between format and the dichotomized *dose-risk estimates* indicated that males responded similarly for both image formats and females for the AN and plain formats. Subsequent 2 × 2 Chi-square analyses that combined formats with similar responses showed that more males assigned to the AN format had accurate numerical *dose-risk estimates* (46.4%) than those assigned to a visual images (21.7%); *Pearson's* $\chi^2(1, n=97) = 5.91$, $p = .015$. Among females there was no difference; accurate *dose-risk estimates* (56.0%) for the graded and (45.1%) for the AN and plain image formats; *Pearson's* $\chi^2(1, n=163) = 1.64$, $p = .200$. Chi-square results were null for the sample split by OTB rather than gender.

Stronger dose-risk beliefs are indicated when there is a larger positive difference in beliefs between those assigned to a high compared to a low dose. The first three rows of Table VII show mean differences in safety beliefs between low and high dose participants split by OTB. The smallest and largest differences between participants assigned to low and high doses were respectively among plain format participants without and with OTB. Results for the sample split by gender are available from the first author.

The graded format was intended to communicate weaker risk at the lower dose and stronger risk at the higher dose compared to the plain format. The last 2 rows of Table VII show mean safety belief differences between the graded and plain formats for each dose. While there was a trend of weaker beliefs at both time points for low dose non-OTB participants assigned to the graded compared to the plain image and stronger beliefs among those assigned to the higher dose at time 2, only 1 of these 4 differences achieved significance.

3.6. RQ 6

How does format influence recall accuracy of the dose and dose-standard relationship?

The plain and graded format subgroups were merged into a visual image group. All males assigned to the AN format correctly recalled their dose compared to 88.6% for the image formats; *Pearson's* $\chi^2(1, N = 98) = 3.48$, $p = .062$. The percentage of females who correctly

[9]Accurate = risk at 40 is about "*twice the risk*" as 20; inaccurate = "*about the same*", "*more than twice*", or "*a little more*".

recalled their dose did not differ between those getting the AN (91.7%) versus the image formats (95.1%); *Pearson's* $\chi^2(1, N = 163) = 0.80$, $p = .372$.

The percentage of males who correctly recalled their dose as over the drinking water standard did not differ between those getting the AN (85.7%) versus the image formats (85.7%); *Pearson's* $\chi^2(1, N=98) = 0.00$, $p < .001$. A larger percentage of females (93.3%) assigned to the AN format correctly recalled their dose as over the standard than those assigned to an image format (81.6%); *Pearson's* $\chi^2(1, N=163) = 4.34$, $p = .037$.

3.7. RQ 7

How does format influence the appropriateness of safety beliefs and risk mitigation intentions among participants with incorrect dose and dose-standard relationship recall?

We used a two-sided Barnard's Exact Test due to the small sample size. Among the 11 participants who incorrectly recalled their dose as less than 10 ppb (safe compared to health standard), more participants assigned to the image had appropriate *safety beliefs* (unsafe) and *risk mitigation intentions* at both time points than those assigned to the AN format, although results for *risk mitigation intentions* were not significant at time 2 (Table VIII). Among the 23 females who incorrectly recalled their assigned dose as under the standard (safe according to standard), more females assigned to the image appropriately rated their water as unsafe and intended to mitigate risk than those assigned to the AN format at both time points, although results for *risk mitigation intentions* were not significant at time 2 (Table IX). Among the 14 males who incorrectly recalled their dose as under the standard, there were no differences in *safety beliefs* or *risk mitigation intentions* to time 1 or 2.

Among the 18 male and female participants with incorrect dose recall, all 7 participants who incorrectly recalled their dose as more than 50 ppb also recalled their result as over the standard. However, the 11 participants who incorrectly recalled their dose as less than 10 ppb were nearly evenly split; 6 incorrectly recalled their dose as under and 5 correctly recalled their dose as over the standard (incorrect dose recall by standard recall, $p = .002$). Among the 6 who incorrectly recalled their dose as less than 10 ppb and under the standard; 5 were assigned to images. All of these 5 image participants had appropriate *safety beliefs* (unsafe) and *risk mitigation intentions*, while the 1 AN participant rated his or her water as safe and did not intend to mitigate risk (format by safety beliefs or by mitigation intentions, $p = .067$).

4. DISCUSSION

Among all participants, only the plain format prompted stronger risk mitigation intentions than the other formats, and these effects endured over time. It is notable that the plain format exerted effects in the presence of three known influences on risk beliefs; dose(1,5,13) gender,(37) and bias.(39,40) These results are discussed along with findings from moderator defined study subgroups.

4.1. Format and Gender Effects on Intentions

Gender effects for responses to the graded image were pronounced and consistent across multiple analyses. The graded image was only related to stronger risk beliefs and mitigation intentions among females and the plain image was only more effective than the graded image among males. The sole difference between these formats was the graded color for the red portion of the scale. The inconsistent response to the graded shades of red is supported by Cleveland and McGill's proposition that shading and color saturation were the least accurately understood of the ten elementary visual features.(10) Visual salience and gestalt may also explain these effects.

4.2. Visual Salience and Gestalt

During the informal development of the images, most people preferred the graded image, but some commented that the plain image was a stronger message because it was more bold and direct "like a stop sign", while the graded image was more subtle. Industrial ergonomists observe that color, high contrast, and icons make warnings more salient (defined as conspicuous or noticeable).(45) Visual salience has been found to shape information acquisition.(46) Proximity, a gestalt principle of information design, conveys relatedness.(47) The test result, drinking water standard, icon, and *stop drinking* message may have appeared more related when grouped within the plain rather than the graded block of the image. The proximity of multiple evidence-based features within the high contrast plain red "warning" interval of the scale that depicted the spatial configuration of the test result and drinking water standard on an uncluttered background may explain the effectiveness of this image over the other formats. Decreased visual salience and perceived relatedness for the graded image may have attenuated the intended message among males. The proposed ability of females to better attend to details than males(36) may explain why gradations were less of a barrier for females. Gender may moderate responses to visual features.

We hypothesized that the graded image would convey weaker and stronger risk respectively for the low and high dose compared to the plain image. Only findings among females or non-OTB participants hinted at this.[10] Adjustments to the color and gradations could potentially increase visual salience and strengthen the intended dose-risk message.

4.3. Format Effects Via Beliefs and Emotions Over Time

Image compared to AN effects were (1) substantial, (2) related to stronger risk beliefs and intentions to mitigate risk, and (3) endured over time with minimal attenuation. Color and the vertical scale were used to convey the *safety* meaning of test results, and the icon and *stop drinking* message to convey the recommended *risk mitigation* response. Time 1 findings were in line with the intended meaning - the plain image format as compared to the AN format directly and indirectly increased intentions to mitigate risk, and most of the indirect effects on these intentions were transmitted via beliefs that water was unsafe. However at time 2, indirect effects were dispersed among specific (likelihood and severity) and global beliefs (safety and serious problem). Format effects on global beliefs were essentially all as direct effects at time 1, but roughly half as indirect effects through specific health risk beliefs at time 2. Findings showed a pattern similar for females assigned to the graded format. These patterns of more distributed indirect effects at time 2 suggest that beliefs became consolidated and integrated over time from an initial gist of water safety toward a gist of integrated safety and health risk. The nature of gist is to knit overarching principles with problem solving,(22) thus the principle of safety as related to health risk may explain this shift over time. Visual salience may have fostered the consolidation process.

The pattern differed among males for the plain relative to the graded format; most format effects on mitigation were direct effects. Since the only difference between these formats was the plain versus graded red block, the more dominant direct effects could be due to the increased visual salience of the *stop drinking* message and icon.

Format effects on the durability of severity beliefs and distress over time add to the above path analysis findings. Severity grew stronger over time among more participants assigned to the images and weaker among more participants assigned to the AN format even though

[10]These trends are consistent with findings that females have greater capacity than males to attend to details, and with findings reported later in the manuscript that non-OTB participants may have responded heuristically to color (in this case the graded color).

images were not designed to convey severity. Mean beliefs across *Δ time subgroups* indicated a consistent trend that stronger distress relative to severity at time one resulted in stronger severity at time 2 and weaker distress relative to severity at time 1 resulted in weaker severity at time 2. Mean severity beliefs were essentially reversed at time 2 time rather than simply trending to the mean. Results suggest that distress at time 1 may have influenced the change within subjects toward time 2 severity beliefs that were more aligned with distress. The combination of visual salience and the warning message of red may have promoted the stronger time 1 emotional response (distress) that shaped stronger severity beliefs at time two.[11] These findings suggest that emotion may have played a role in facilitating the consolidation of gist to integrate health risk beliefs with safety beliefs. This is in keeping with the role of emotion in developing gist,(27) of affect as tightly integrated with cognition,(28) and findings that emotion is especially integrated with beliefs about health consequences.(48)

4.4. Accuracy of Recall for Facts and Beliefs

Despite the relationship of images to more appropriate beliefs and intentions, fact-based recall was more accurate among participants assigned to the AN format than images. Males better recalled their dose amount (and had more accurate dose-risk estimates) and females better recalled their dose as over the drinking water standard. Chua(30) also found that participants assigned to numerical information had better recall of numerical facts than those assigned to graphs. Presentation and task interact to influence recall(49) and cues can prompt verbatim recall.(50) The use of text and numbers in the AN format and the survey question to respectively convey and then assess recall of both dose amount and relationship to the standard may have interacted to enhance verbatim recall for the AN format.

The enhanced recall and numerical accuracy of dose-risk estimates for the AN format did not result in more appropriate safety beliefs or intentions to drink less water. This supports the observation that formats related to more accurate knowledge may differ from those related to appropriate global beliefs or behavior.(29) Images designed to convey the meaning of risk may promote enduring and meaningful risk beliefs (gist) that are more applicable for making decisions than accurate but less-meaningful facts.(3)

The most provocative findings in this study pertained to the few participants that inaccurately recalled either their dose as less than 10 or their test result as less than the standard. Among these participants, and even among those who incorrectly recalled both, more of those assigned to images had appropriate safety beliefs and mitigation intentions than those assigned to the AN format. These findings are speculative due to the small sample, however, they suggest that images may have the capacity to convey accurate gist-based meaning independent of recalled facts. This is consistent with the fuzzy trace theory proposition and related findings that cognition results in parallel but separate memory traces for verbatim and gist-based knowledge.(21)

Findings of the AN format supporting accurate recall of facts, the image versions supporting enhanced development of meaningful gist, and correct recall of meaning among participants who incorrectly recalled the facts are all congruent with fuzzy trace theory.(21,50)

4.5. Images, Biased Beliefs and Safety Threshold Bias

Images only prompted meaningful risk reduction intentions among participants with OTB. Participants without OTB may have applied the drinking water standard as the dominant

[11] In this study, the severity of health risks had a stronger correlation with emotion than the likelihood of health risks.

heuristic[12] for choosing mitigation intentions, potentially weakening the impact of the meaning conveyed by image features.

The plain image appeared to prompt a safety standard *threshold effect* among individuals without OTB. Among these participants, safety beliefs were very similar for those assigned to the low and high dose, perhaps because both were in the same “above threshold” color block of the scale. This was not evident for non-OTB participants assigned to the AN or graded formats, perhaps because these formats conveyed numerical or visual dose-risk relationships. These findings suggest heuristic responses to number or color formats among non-OTB participants.

Conversely, the plain image prompted the strongest dose-risk beliefs among OTB individuals. Individuals with OTB may have deliberated more in selecting a personal safety threshold that was less rigorous than the drinking water standard. This choice may have been supported by dose-risk beliefs that did not strongly adhere to the safety standard. For example, if they believed the difference in associated risk between 9 and 11 ppb was minimal at a safety standard of 10, they may have felt more comfortable nudging their personal safety standard upwards. The higher contrast in the plain red block of the vertical scale may have fostered stronger configurational knowledge because it clearly depicted the magnitude and direction of the spatial relationships for the low or the high dose relative to the safety standard. One would expect this influence to be stronger at time 1 when visual features were present as was indicated by study findings. Findings suggest that the plain image may have been used differently by participants with and without OTB. For non-OTB participants, the dominant visual feature may have been the symbolic meaning of color (red = danger) and the stop drinking icon. This suggests a heuristic application of these symbols based on a learned convention. For OTB participants, the dominant visual feature may have been the vertical scale. The spatial magnitude of the low or high dose as near or far from the standard may have supported their beliefs that results near the standard are of less concern than those higher up the scale.

Threshold bias may have generated strong effects because it included elements of both an independent and dependent variable. The weak positive relationship between dose and the personal safety threshold suggested that OTB was partially induced by threat thus consistent with defensive denial studies.(6,7,8) Inherent optimism(40) and/or skepticism of health advice(33) may explain how this variable operated as an independent variable. OTB had more explanatory value as an independent variable because it illustrated that images were more effective among biased individuals and suggested that the same image may be used differently based on the presence or absence of bias and, perhaps, on information processing.

5. LIMITATIONS

Information content was only strictly controlled between the plain and graded formats and thus constrained our ability to draw firm conclusions about differences between AN and image information. Data collection procedures of study participation after class may have prompted quick responses that exacerbated gender differences if the additional visual detail in the graded image was more easily apprehended by females. Quick judgments may have favored less deliberate and analytic processing. In practice, some people may briefly scan test results, while others may study them more carefully. This study approximates the former but not the latter condition. College students’ stronger numeracy and analytical skills and typical residence in dwellings served by municipal water likely influenced their responses to

[12]A heuristic is a decision rule or learned convention that replaces deliberate systematic information processing.(56)

the study information. Some analysis was across moderator defined sub-groups and thus increased the chance of generating spurious findings. Follow-up work is needed to further test these results.

6. IMPLICATIONS FOR PRACTICE AND FURTHER STUDY

The exploratory nature of this study and the student sample constrain our ability to make strong recommendations for practice. Results suggest that images designed to convey the meaning of risk information can close the gap between the intended and imparted meaning of environmental health risk information. The enhanced effectiveness of images among participants with biased beliefs is especially promising since these individuals may be less likely to attend and respond to risk information. The utility of fuzzy trace theory for explaining the role of gist-based beliefs compared to accurate fact-based knowledge, the consolidation of beliefs over time, and the role of emotion suggest its value for risk communication research. Findings hint at the potential ability of images to convey accurate meaning independent of detailed facts. Even when individuals have an inaccurate understanding of the detailed facts, they can accurately understand and act on the meaning extracted from well designed information. Differences for the effects of plain and graded images and the more durable beliefs among participants assigned to images highlight the importance of visual salience for graphical displays. The potential power of visual images indicates a need to develop ethical guidelines to guide the appropriate use of image features. For example, when safety standards or evidence-based recommendations for mitigation do not exist, the use of symbols to concretely convey safety or recommended actions may communicate more certainty about a given risk than warranted.

This exploratory study prompts many suggestions for further research. The personal relevance of test result information may influence the meaning derived from visual information, therefore, the study should be replicated among a target sample of private well owners. Studies should explore responses to images showing benign results since images should accurately convey the meaning of both safe and unsafe results. Although practical applications include integrating multiple image features in a single image, meaning-based beliefs may be differentially influenced by these features. Experimental work needs to explore how these features differentially shape beliefs and behavior. Gender differences in apprehending and remembering visual detail and relationships between image features, bias, information processing, and memory need further exploration. Future studies should be designed to examine how images and image features shape the development of gist, and the influence of emotion and visual salience on that process. Numeracy influences health risk information processing;(51) inclusion as a potential moderator could provide insights into its role in processing visual risk information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a post-doctoral fellowship through UW-Madison’s Computation and Informatics in Biology and Medicine program funded by the National Library of Medicine (5T15LM007359). We thank the Brennan Health Systems Lab for support in developing this research, Lester Dore for assistance with the images, and Roger Brown for advice on the path analysis. All of these individuals are affiliated with the UW-Madison School of Nursing.

REFERENCES

1. Severtson DJ, Baumann LC, Brown RL. Applying a health behavior theory to explore the influence of information and experience on arsenic risk representations, policy beliefs, and protective behavior. Risk Analysis 2006;26(2):356–372.
2. Bostrom A, Anselin L, Farris J. Visualizing Seismic Risk and Uncertainty. A Review of Related Research. Annals of the New York Academy of Sciences 2008;1128(1):29–40. [PubMed: 18469212]
3. Lipkus IM. Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations. Medical Decision Making 2007;27(5):696–713. [PubMed: 17873259]
4. Johnson B. Public Views on Drinking Water Standards as Risk Indicators. Risk Analysis 2008;28(6):1515–1530. [PubMed: 18793283]
5. Weinstein, N.; Sandman, PM.; Roberts, NE. Report to U.S. EPA from Environmental Communication Research Program. Cook College, Rutgers University; New Brunswick, NJ: 1989. Report nr EPA-230-08-89-064
6. Ditto, PH.; Croyle, RT. Understanding the impact of risk factor test results: Insights from a basic research program. In: Croyle, RT., editor. Psychological effects of screening for disease prevention and detection. Oxford University Press; New York: 1995. p. 144-181.
7. Croyle RT, Barger SD, Loftus EF, Sun YC, Hart M, Gettig J. How well do people recall risk factor test results? Accuracy and bias among cholesterol screening participants. Health Psychology 2006;25(3):425–432. [PubMed: 16719615]
8. Glanz K, Brekke M, Hoffman E, Admire J, McComas K, Mullis R. Patient reactions to nutrition education for cholesterol reduction. American Journal of Preventive Medicine 1990;6:311–317. [PubMed: 2076297]
9. Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences 2001;24:97–185.
10. Lohse GL. A cognitive model for understanding graphical perception. Human-Computer Interaction 1993;8:353–388.
11. Cleveland WS, McGill R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association 1984;79(387): 531–554.
12. Golledge, RG.; Stimson, RJ. Analytical Behavior Geography. Croom Helm; New York: 1987. Learning and spatial behavior; p. 84-105.
13. Sandman PM, Weinstein ND, Miller P. High Risk or Low: How Location on a “Risk Ladder” Affects Perceived Risk. Risk Analysis 1994;14(1):35–45. [PubMed: 8146401]
14. Harrower M, Brewer CA. ColorBrewer.org: An online tool for selecting colour schemes for maps. The Cartograph Society 2003;40(1):27–37.
15. MacEachren, AM. How Maps Work: Representation, Visualization, and Design. The Guilford Press; New York: 2004. How maps are imbued with meaning; p. 213-215.
16. Griffith LJ, Leonard SD. Association of colors with warning signal words. International Journal of Industrial Ergonomics 1997;20:317–325.
17. Murray LA, Magurno AB, Glover BL, Wogalter MS. Prohibitive pictorials:: Evaluations of different circle-slash negation symbols. International Journal of Industrial Ergonomics 1998;22(6): 473–482.
18. Robinson, AH.; Sale, RD.; Morrison, JL.; Muehrcke, PC. Elements of Cartography. Wiley; New York: 1984.
19. Young SL, Wogalter MS. Comprehension and memory of instruction manual warnings: Conspicuous print and pictorial icons. Human Factors 1990;32:637–649.
20. Kohler, W. Gestault psychology. Liveright; New York: 1947.
21. Reyna VF, Brainerd CF. Fuzzy-trace theory: An interim synthesis. Learning and Individual Differences 1995;7(1):1–75.
22. Reyna VF, Brainerd CF. Fuzzy-trace theory and children’s acquisition of mathematical and scientific concepts. Learning and Individual Differences 1991;3(1):27–59.

23. Zhang J, Norman DA. Representations in distributed cognitive tasks. Cognitive Science 1994;18(1):87–122.
24. Weinstein ND. The precaution adoption process. Health Psychology 1988;7:355–386. [PubMed: 3049068]
25. Wolfe CR. Cognitive technologies for gist processing. Behavior Research Methods 2006;38(2): 183–189. [PubMed: 16956092]
26. Valdez P, Mehrabian A. Effects of color on emotions. Journal of Experimental Psychology: General 1994;123(4):394–409. [PubMed: 7996122]
27. Rivers SE, Reyna VF, Mills B. Risk Taking Under the Influence: A Fuzzy-Trace Theory of Emotion in Adolescence. Developmental Review 2008;8(1):107–144. [PubMed: 19255597]
28. Slovic P, Finucane ML, Peters E, MacGregor DG. Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. Risk Analysis 2004;24(2):311–322. [PubMed: 15078302]
29. Ancker JS, Senathirajah Y, Kukafka R, Starren JB. Design Features of Graphs in Health Risk Communication: A Systematic Review. J Am Med Inform Assoc 2006;13(6):608–618. [PubMed: 16929039]
30. Chua HF, Yates JK, Shah P. Risk avoidance: Graphs versus numbers. Memory & Cognition 2006;34(2):399–410.
31. Leonard SD. Does color of warnings affect risk perception? International Journal of Industrial Ergonomics 1999;23:499–504.
32. Strecher, VJ.; Kreuter, MW. The psychosocial and behavioral impact of health risk appraisals. In: Croyle, RT., editor. Psychological effects of screening for disease prevention and detection. Oxford University Press; New York: 1995. p. 126-143.
33. Frankel S, Davison C, Smith GD. Lay epidemiology and the rationality of responses to health education. British Journal of General Practice 1991;41(351):428–430. [PubMed: 1777300]
34. Lipsey, MW. Design sensitivity: Statistical power for experimental research. Sage Publications; Newberry Park, CA: 1990.
35. Baron RM, Kenny DA. The moderator-mediator variable distinction in social-psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 1986;51(6):1173–1182. [PubMed: 3806354]
36. Levy-Meyers, J. Gender differences in cortical organization: Social and biochemical antecedents and advertising consequences. In: Clark, EM.; Brock, TC.; Stewart, DW., editors. Attention, Attitude, and Affect in Response to Advertising. Lawrence Erlbaum; 2004.
37. Slovic P. Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. Risk Analysis 1999;19(4):689–701. [PubMed: 10765431]
38. Halpern, DF.; Collaer, ML. Sex differences in visuospatial abilities: More than meets the eye. In: Shah, P.; Miyake, A., editors. The Cambridge handbook of visuospatial thinking. Cambridge University Press; New York: 2005. p. 170-212.
39. Weinstein ND, Klein WM. Resistance of personal risk perceptions to debiasing interventions. Health Psychology 1995;14(2):132–140. [PubMed: 7789348]
40. Schwarzer R. Optimism, vulnerability, and self-beliefs as health-related cognitions: A systematic overview. Psychology and Health 1994;9:161–180.
41. Barrett A, et al. Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction, and number needed to treat. Canadian Medical Association Journal 2004;171(4): 353–358. [PubMed: 15313996]
42. UBC Clinical Significance Calculator. Vol. Volume 2008. University of British Columbia;
43. Muthen, LK.; Muthen, B. O. Mplus. 4.21. Muthen & Muthen; 2007.
44. Muthen, LK.; Muthen, B. O. Mplus User's Guide. Muthen & Muthen; Los Angeles, CA: 2007.
45. Wogalter MS, Conzola VC, Smith-Jackson TL. Research-based guidelines for warning design and evaluation. Applied Ergonomics 2002;33(3):219–230. [PubMed: 12164506]
46. Jarvenpaa SL. Graphic displays in decision making: The visual salience effect. Journal of Behavioral Decision Making 1990;3:247–262.

47. Ware, C. Information Visualization: Perception for Design. Morgan Kaufmann; San Francisco: 2000.
48. Severtson DJ, Baumann LC, Brown RL. Applying a the common sense model to understand representations of arsenic contaminated well water. Journal of Health Communication 2008;13(6): 523–554. [PubMed: 18726810]
49. Washburne JN. An experimental study of various graphic, tabular, and textual methods of presenting quantitative material. Journal of Educational Psychology 1927;18(6):361–376.
50. Reyna VF, Brainerd CJ. Fuzzy-trace theory and memory development. Developmental Review 2004;24(4):396–439.
51. Fagerlin A, Ubel PA, Smith DM, Zikmund-Fisher BJ. Making numbers matter: Present and future research in risk communication. American Journal of Health Behavior 2007;31:S47–S56. [PubMed: 17931136]
52. Shah, PR.; Freedman, EG.; Vekiri, I. The comprehension of quantitative information in graphical displays. In: Shah, PR.; Miyake, A., editors. The Cambridge handbook of visuospatial thinking. Cambridge University press; New York: 2005. p. 426-476.
53. Lipkus IM, Hollands JG. The visual communication of risk. Journal of National Cancer Institute 1999;25:149–162.
54. MacEachren, AM. How maps work: representation, visualization, and design. Guilford; New York: 2004.
55. Trinder, L.; Reynolds, S. Introduction: The context of evidenced-based practice. In: Trinder, L., editor. Evidenced-based practice: A critical appraisal. Wiley-Blackwell; New York: 2000. p. 1-16.
56. Chaiken, S.; Liberman, A.; Eagly, AH. Heuristic and systematic information processing within and beyond the persuasion context. In: Uleman, JS.; Bargh, JA., editors. Unintended thought. Guilford; New York: 1989. p. 212-252.

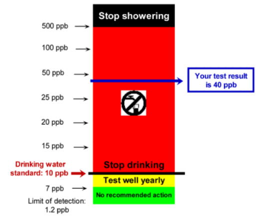

Figure 1.
Plain Red Image

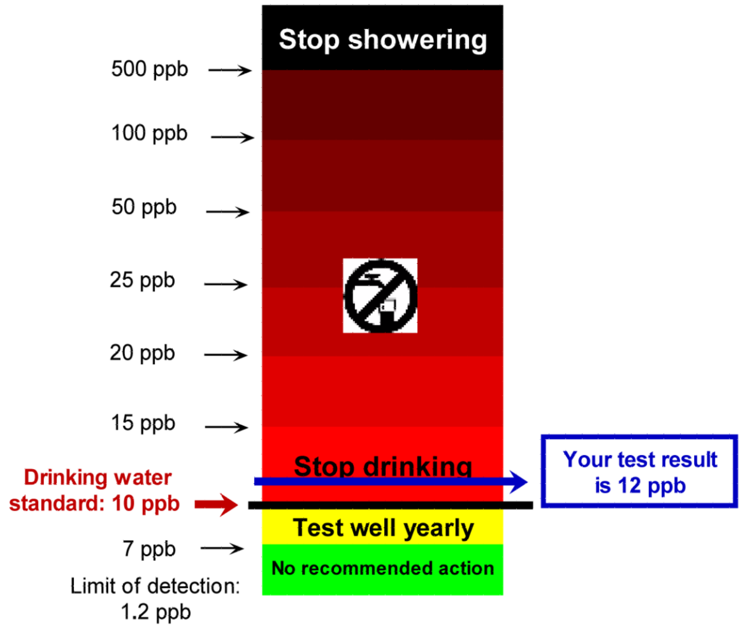

Figure 2.
Graded Red Image

Test	Result	Units	LOD	LOD	Method	Analyzed	Analyst
Manganese, dissolved	12*	ppb	1.100	3.800	3113B	2/8/06	CB

Figure 3.
Alphanumeric Format

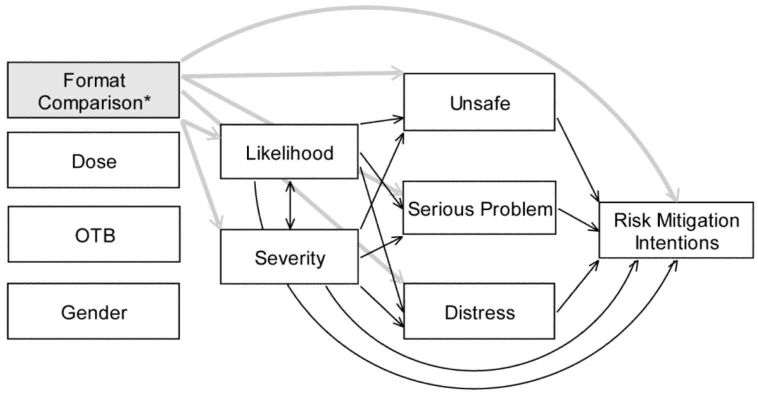

Figure 4.
Conceptual Model: Format Effects on Risk Mitigation Intentions Via Specific and Global Beliefs and Emotion
*Only paths for format effects are depicted here, although dose, OTB, and gender were included in the model. Direct format effects are depicted with thick gray model path, values for these effects are in Table IV.

Table I

4-way ANOVA: Format, Dose, Gender and Optimistic Threshold Bias on Risk Mitigation at Time 1 and Time 2

		Time 1[a]		Time 2[b]	
Source	**df**	**F**	***p***	**F**	***p***
Format	2	5.39	.005	4.70	.010
Dose	1	24.55	<.001	40.93	<.000
Gender	1	4.17	.042	3.17	.076
Optimistic Threshold Bias (OTB)	1	137.91	<.000	140.76	<.000
Format * Gender	1	2.77	.065	1.28	.281
Dose * OTB	1	2.93	.088	1.34	.248
Gender * OTB	1	8.50	.004	8.27	.004

[a]Adjusted R^2 = .435;

[b]Adjusted R^2 = .442

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table II

Post hoc Comparisons* for Format

Format Comparisons	Time 1 Δ M	Time 1 *SE*	Time 1 *p*	Time 2 Δ M	Time 2 *SE*	Time 2 *p*
All (n = 261)						
Alphanumeric-Plain	0.54	.145	.001	0.47	.142	.003
Alphanumeric-Graded	0.24	.147	.289	0.22	.144	.343
Plain-Graded	−0.30	.147	.120	−0.26	.144	.214
Females (n = 153)						
Alphanumeric-Plain	0.54	.178	.009	0.46	.183	.038
Alphanumeric-Graded	0.61	.181	.003	0.50	.186	.022
Plain-Graded	0.08	.186	.968	0.04	.191	.995
Males (n = 98)						
Alphanumeric-Plain	0.59	.253	.067	0.54	.226	.054
Alphanumeric-Graded	−0.27	.257	.653	−0.15	.229	.887
Plain-Graded	−0.86	.240	.002	−0.69	.215	.0

* With Sidak Corrections for Multiple Comparisons

Table III

Relative Risk Reduction (RRR) and Chi-square for Risk Mitigation Intentions

STB	Gender	[O] C-T n	Comparison Group	Treatment Group	χ^2	Time 1 RRR	χ^2	Time 2 RRR
No	Male	8-15	Alphanumeric	Plain	-	0%	[0.56]	[7%]
No	Male	8-7	Alphanumeric	Graded	-	0%	-	0%
No	Male	7-15	Graded	Plain	-	0%	0.49	8%
Yes	**Male**	**20-21**	**Alphanumeric**	**Plain**	**2.97** $^{+}$	**77%**	**2.11**	**73%**
Yes	Male	20-27	Alphanumeric	Graded	0.16	17%	0	0%
Yes	**Male**	**27-20**	**Graded**	**Pain**	**2.12**	**51%**	**2.56**	**73%**
No	Female	27-19	Alphanumeric	Plain	1.47	8%	1.05	12%
No	Female	27-23	Alphanumeric	Graded	0.21	3%	0.44	7%
No	Female	23-19	Graded	Plain	0.85	4%	0.19	4%
Yes	**Female**	**33-34**	**Alphanumeric**	**Plain**	**2.57**	**37%**	**1.19**	**27%**
Yes	**Female**	**33-27**	**Alphanumeric**	**Graded**	**4.41** *	**50%**	**1.26**	**29%**
Yes	Female	27-34	Graded	Plain	[0.40]	9%	[0.01]	2%

[O] Number of participants for the Comparison group (C) and Treatment group (T)

Values with Relative Risk Reduction Values ≥ 25% are bolded

[Values in brackets showed higher rates of mitigation for the control group]

$^{+}p < .10$;

$^{*}p < .05$

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table IV

Total, Direct, and Total Indirect Effects[1] Generated by Format

	AN - Plain		AN - Graded Female		Graded - Plain Male	
Format Effects[2]	**TIME 1 CFI = .903 SRMR = .039 n = 177**	**TIME 2 CFI = .908 SRMR = .037 n = 177**	**TIME 1 CFI = 0.893 SRMR =.048 n=110**	**TIME 2 CFI = .907 SRMR =.042 n=110**	**TIME 1 CFI =.867 SRMR =.053 n=62**	**TIME 2 CFI =.913 SRMR =.036 n=62**
On Risk Mitigation	R^2 = .68	R^2 = .70	R^2 = .67	R^2 = .68	R^2 = .70	R^2 = .75
Total effects	.49***	.45***	.51***	.43**	.40*	.30+
Direct effects	.19*	.20*	.08	.14	.30*	.24+
Total indirect effects	.30***	.25***	.43***	.28**	.10	.06
∞Via likelihood	.02	.09	.02	.07	.04	.04
∞Via severity	−.02	.08	.05	.07	.03	.04
∞Via safety belief	.21	.07	.29	.07	.06	0
∞Via serious problem	.05	.10	.04	.08	.03	.08
∞Via distress	.04	.02	.07	.07	−.02	−.01
On Safety Beliefs	R^2 = .58	R^2 = .65	R^2 = .56	R^2 = .67	R^2 = .561	R^2 = .662
Total effects	.63***	.40***	.73***	.51***	.25	−.02
Direct effects	.63***	.20*	.67***	.38***	.18	−.17
Total indirect effects	0	.20**	.06	.13	.07	.15
Via severity	−.01	.09*	.04	.03	.03	.11
Via likelihood	.01	.11*	.02	.10	.04	.04
On Serious Problem	R^2 = .60	R^2 = .70	R^2 = .61	R^2 = .66	R^2 = .549	R^2 = .771
Total Effects	.36***	.37***	0.47***	.38*	.09	.22
Total direct effects	.37***	.14+	.37***	.21*	0	.05
Total indirect effects	−.01	.23**	.10	.17	.09	.17
Via severity	−.02	.13**	.08	.09	.05	.12

Format Effects[2]	AN - Plain		AN - Graded Female		Graded - Plain Male	
	TIME 1 CFI = .903 SRMR = .039 n = 177	**TIME 2 CFI = .908 SRMR = .037 n = 177**	**TIME 1 CFI = 0.893 SRMR =.048 n=110**	**TIME 2 CFI = .907 SRMR =.042 n=110**	**TIME 1 CFI =.867 SRMR =.053 n=62**	**TIME 2 CFI =.913 SRMR =.036 n=62**
Via likelihood	.01	.10*	.02	.08	.04	.05
On Distress	R^2 = .54	R^2 = .55	R^2 = .47	R^2 = .54	R^2 = .55	R^2 = .54
Total Effects	.44***	.44***	.38**	.45**	.26	.11
Direct effects	.45***	.25*	.29*	.30*	.17	−.03
Total indirect effects	−.01	.19**	.09	.15	.08	.14
Via severity	−.02	.12*	.07	.09	.05	.10
Via likelihood	.01	.07^{+}	.02	.06	.03	.04

[1] Y-standardized effects

[2] Format effects are defined as follows: *Direct Effects* = path from format to the designated dependent variable, *Total Indirect Effects* = sum of all paths from format to the designated variable that occur indirectly through another variable, *Total Effects* = Direct Effects + Total Indirect Effects; *Total Indirect Effects Via* (*likelihood, severity*…..) = Sum of all indirect effects that occur through that variable.

$^{\infty}$ p-values are not available for these summed indirect effects via each of these variables

$^{+}$ $p < .10$,

* $p < .05$,

** $p < .01$,

*** $p < .001$

Table V

Proportion of Change over Time with X^2 for *Severity Δ Time Subgroups* Between Format Pairs

Δ Time Subgroups Based on Severity	**Severity** All[1] Pair n =177		All[2] Pair n =172	
	AN	**Plain**	**AN**	**Graded**
Stronger	11.4% n=10	25.8% n=23	11.4% n=10	29.8% n=25
No change	44.3% n=39	47.2% n=42	44.3% n=39	33.3% n=28
Weaker	44.3% n=39	27.0% n=24	44.3% n=39	36.9% n=31
Total	88 100.0%	89 100.0%	88 100.0%	84 100.0%

[1] $X^2 = 8.80$, p=.012

[2] $X^2 = 9.06$, p=.011

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table VI

Time 1 and Time 2 Mean Severity and Distress and Δ Time t-tests Across *Severity Δ Time Subgroups*

	All Participants Excluding Males with Graded Format (n=227)					
Δ Time	**Time 1**		**Time 2**		**T1 T2 t-test**	
Subgroups Based on Severity	**Severity** Mean / *SD* / n	**Distress** Mean / *SD* / n	**Severity** Mean / *SD* / n	**Distress** Mean / *SD* / n	**Severity** *t* / *p-value* / df	**Distress** *t* / *p-value* / df
Stronger	3.11 / *1.255* / 47	3.60 / *1.014* / 47	4.34 / *1.185* / 47	3.68 / *0.783* / 47	15.11 / *< .001* / 46	0.85 / *0.400* / 46
No change	3.92 / *1.172* / 98	3.75 / *1.124* / 98	3.92 / *1.172* / 98	3.71 / *0.984* / 98	0	−0.46 / *0.650* / 97
Weaker	4.61 / *1.003* / 82	3.74 / *1.087* / 82	3.30 / *1.108* / 82	3.55 / *0.971* / 82	−22.00 / *< .001* / 81	−2.37 / *0.020* / 81
All groups	4.00 / *1.255* / 227	3.71 / *1.086* / 227	3.78 / *1.213* / 227	3.65 / *0.940* / 227	−3.07 / *0.002* / 226	−1.41 / *0.159* / 226

Table VII

Mean Difference for Safety Beliefs: 12 versus 40 ppb and Graded versus Plain Image

Safety belief differences	No OTB Time 1 Mean *SE*	No OTB Time 2 Mean *SE*	OTB Time 1 Mean *SE*	OTB Time 2 Mean *SE*
40 ppb minus 12 ppb				
AN n = 35[NB]; n = 53[B]	0.81* .302	0.47 .295	0.85** .272	0.80** .263
Plain n = 34[NB]; n = 55[B]	0.15 .265	0.06 .226	1.13*** .247	0.88*** .251
Graded, n = 30[NB]; n = 54[B]	0.67* .259	0.67** .220	0.37 .316	0.69** .257
Graded minus Plain				
12 ppb n = 35[NB]; n = 52[B]	−0.32 .240	−0.50* .194	0.22 .296	0.22 .279
40 ppb n = 29[NB]; n = 57[B]	0.20 .290	0.22 .257	−0.54 .272	0.03 .230

NB = No STB

B = STB

*Significance based on Independent T-test; $p < .05$;

** $p < .01$;

*** $p < .001$

Table VIII

Frequencies for Dichotomous Safety Beliefs and Mitigation Intentions among Alphanumeric and Image Participants who Incorrectly Recalled Dose < 10 ppb

	Time 1[1]		Time 2[2]	
	AN	**Image**	**AN**	**Image**
Safe	2	0	2	1
Unsafe	1	8	1	7
No mitigate	2	0	1	1
Mitigate	1	8	2	7

[1] Barnard's Exact Test: $p = .020$ for safety beliefs and mitigation intentions

[2] Barnard's Exact Test: $p = .117$ for safety beliefs; $p = .575$ for mitigation intentions

Table IX

Frequencies for Dichotomous Safety Beliefs and Mitigation Intentions among Alphanumeric and Image Female Participants with Incorrect Dose-Standard Relationship Recall

	Time 1[1]		Time 2[2]	
	AN	**Image**	**AN**	**Image**
Safe	3	3	3	3
Unsafe	1	16	1	16
No mitigate	2	0	1	2
Mitigate	2	19	3	17

[1]Barnard's Exact Test: $p = .017$ or safety beliefs; $p = .008$ for mitigation intentions

[2]Barnard's Exact Test: $p = .017$ for safety beliefs; $p = .573$ for mitigation intentions