COMPREHENSION OF HEALTH RISK PROBABILITIES: THE ROLES OF AGE, NUMERACY, FORMAT, AND MENTAL REPRESENTATION

A Dissertation
Presented to
The Academic Faculty

by

Cara Bailey Fausset

In Partial Fulfillment
of the Requirements for the Degree
Ph.D. in the
School of School of Psychology

Georgia Institute of Technology
August 2012

COMPREHENSION OF HEALTH RISK PROBABILITIES: THE ROLES OF AGE, NUMERACY, FORMAT, AND MENTAL REPRESENTATION

Approved by:

Dr. Wendy A. Rogers, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Bruce N. Walker
School of Psychology
Georgia Institute of Technology

Dr. Ute M. Fischer
School of School of Literature, Communication, and Culture
Georgia Institute of Technology

Dr. Francis T. Durso
School of Psychology
Georgia Institute of Technology

Dr. Katharina V. Echt
School of Medicine
Emory University
Altanta VA Rehabilitation R&D Center

Date Approved: June 25, 2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Psychologists, healthcare practitioners, public health workers, and others obtaining informed consent must understand how to communicate risk information effectively to ensure comprehension. Probabilities, an essential dimension of risk communication, can be presented in various formats including frequencies (e.g., 1 in 10), percentages (e.g., 10%), or verbal phrases (e.g., unlikely); the literature is mixed concerning which format best supports comprehension. Additionally, it is not well understood how people who vary in their level of numeracy, or ability "to comprehend, use and extract meaning from numbers" (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008, p. 262) understand those probabilities.

Evidence suggests that people, especially older adults, have difficulty comprehending and using probabilities (e.g. Gigerenzer, Gaissmaier, Kurz-Milke, Schwartz, & Woloshin, 2007; Lipkus, Samsa, & Rimer, 2001). People who are low in numeracy represent an at-risk population: If they do not understand probabilities used to express the likelihood of risk, they cannot make informed decisions (Finucane & Gullion, 2010; Finucane, Mertz, Slovic, & Schmidt, 2005; Finucane, Slovic, Hibbard, Peters, Mertz, & MacGregor, 2002; Weinstein, 1999).

Additionally, it has been posited that higher numerate people have more precise mental representations of quantitative values, which can facilitate comprehension (Peters, Slovic, Vastfjall, & Mertz, 2008). The relationship between numeracy, age, precision of mental representations and comprehension is unknown.

The details of how and what people understand when presented with a probability are not well understood, nor how the factors of format and numeracy influence comprehension and mental representations of probabilities for younger and older adults. The goal of the present three-phase within-participant study was to understand how these factors interact and influence comprehension of health risk probabilities.

The first phase of the study used multiple measures to assess participants' comprehension of health risk probabilities expressed as frequencies, percents, and words. The first measure of comprehension was defined as accuracy on comprehension questions about the probabilities. Participants' descriptions of the probability risk expressions were also examined. This approach indirectly assessed the mental representation of probabilities by examining if participants acknowledged the health risk probabilities and how accurately they expressed the probabilities. Additionally, participants' recall of probabilities on a delayed cued recall test was examined. There was a significant effect of format and age such that percent was optimal for supporting both younger and older adults' comprehension and for immediate and delayed representations.

The aim of the second phase was to obtain insight into mental representations of probabilities. Using a magnitude comparison task, format was manipulated and the effects on younger and older adults' accuracy and precision of numerical values were investigated. Percent format led to the highest accuracy; frequency and words were lower. Additionally, there was a distance effect for words: Both younger and older adults were faster on average to compare distant trials than near trials. This pattern is consistent with the idea that verbal expressions are represented spatially on a mental number line. There were no age-related differences in distance effect slopes by age.

The third phase combined the results of the first two phases to determine how format, numeracy, age, mental representation and comprehension were related. Numeracy was strongly related to older adults' accuracy on comprehension questions, but counter to the framework of numeracy (Lipkus & Peters, 2009) and the findings of Peters et al. (2008), no significant correlations between numeracy and precision of mental representation were identified. Additionally, numeracy was a significant predictor of comprehension.

Overall, the results of this research clearly indicated that comprehension and mental representation of health risk probabilities are influenced by format, age, and numeracy. To best support comprehension and comparison of health risk probabilities for younger adults and healthy older adults with varying numeracy, percent format should be used.

CHAPTER 1: INTRODUCTION

Risk communication is an important facet in our lives from the mundane, "There is a 30% chance of rain tomorrow," to the life-threatening, "2 in 100 patients die from complications from this surgery." People must understand these probabilistic expressions of risk to make informed decisions. Weinstein (1999) provided core dimensions that are required for understanding a risk: Comprehending the concept of probability is one of those core attributes and was the focus of this dissertation.

Evidence suggests that people, especially older adults, have difficulty comprehending and using quantitative information, including probabilities (Gigerenzer, Gaissmaier, Kurz-Milke, Schwartz, & Woloshin, 2007; Gigerenzer, Hertwig, van den Broek, Fasolo, & Katsikopoulous, 2005; Kutner, Greenberg, & Baer, 2005; Lipkus, Samsa, & Rimer, 2001; Paulos, 1989; Schwartz, Woloshin, Black, & Welch, 1997). People who are low in numeracy, which is the ability to understand quantitative or numerical information, represent an at-risk population: If they do not understand numerical, probabilistic, graphical, or statistical information, they cannot make informed decisions (Finucane & Gullion, 2010; Finucane, Mertz, Slovic, & Schmidt, 2005; Finucane, Slovic, Hibbard, Peters, Mertz, & MacGregor, 2002; Weinstein, 1999).

Moreover, numeracy level and the format (e.g., frequency, percent, words) used to express the risk probability interact (Peters et al., 2006). Higher numerate people were less influenced by the format in which a probability was presented; that is, risks were rated equivalently regardless of format. Lower numerate individuals, on the other hand, were influenced by the format such that a risk presented as a percent was rated significantly lower than the equivalent risk presented in a frequency format.

Additionally, it has been proposed that numeracy reflects the precision of representation of quantitative values which therefore facilitates comprehension (Peters, Slovic, Vastfjall, & Mertz, 2008).

At this point, however, the details of how people comprehend and mentally represent probabilities are not well understood, nor how the factors of format, age-related differences in abilities, and numeracy influence comprehension and mental representation. The goal of this three-phase study was to improve our understanding of risk communication by systematically investigating how these factors influence comprehension of probabilities.

Understanding Probabilities in Health Risk Communication

Risks are communicated in many domains from using household cleaning products to financial investments to various medications, vaccines, and medical procedures. Comprehension of risk communication is essential for making informed decisions (e.g., Finucane & Gullion, 2010). However, the extant research often assumes numerical information, such as probabilities, is comprehended; the primary dependent variable is the outcome behavior or decision made by the participants (e.g., Schwartz, Woloshin, & Welch, 2005). Although predicting the decisions that people make is important, it cannot be assumed that people understand the information they are given and that they are therefore making informed decisions. In fact, Schapira and colleagues (2008) identified several participants in a focus group study who said “probably” instead of “probability.” “Probably” expresses a notion that an event will most likely happen, whereas “probability” expresses the mathematical notion of degree of the likelihood of an event occurring. These data suggest that some people do not understand the concept of probability.

Thus, a critical gap in the literature is an understanding of the comprehension component of the decision making process with respect to probabilities. A systematic investigation of what people understand and how they mentally represent probabilities presented in different formats as a function of age and numeracy is lacking.

The difficulty lies in defining what it means for a person to understand a risk. Weinstein (1999) provided three basic dimensions or attributes of risk that must be understood: (1) the probability of the risk, (2), the severity of the risk, and (3) ease or difficulty of carrying out actions to reduce the risk. The focus of this dissertation was on the first of these—understanding risk probability—because this dimension requires some minimum level of numeracy, the ability “to comprehend, use and extract meaning from numbers” (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008, p. 262).

Factors that Influence Comprehension of Probabilities

Two recent frameworks have been proposed that illustrate the role of numeracy in decision making and outcome behaviors in a health context (Lipkus & Peters, 2009; Reyna, Nelson, Han, & Dieckmann, 2009). Comprehension was central to the framework proposed by Lipkus and Peters shown in Figure 1, whereas comprehension was only implied in the Reyna et al. framework.

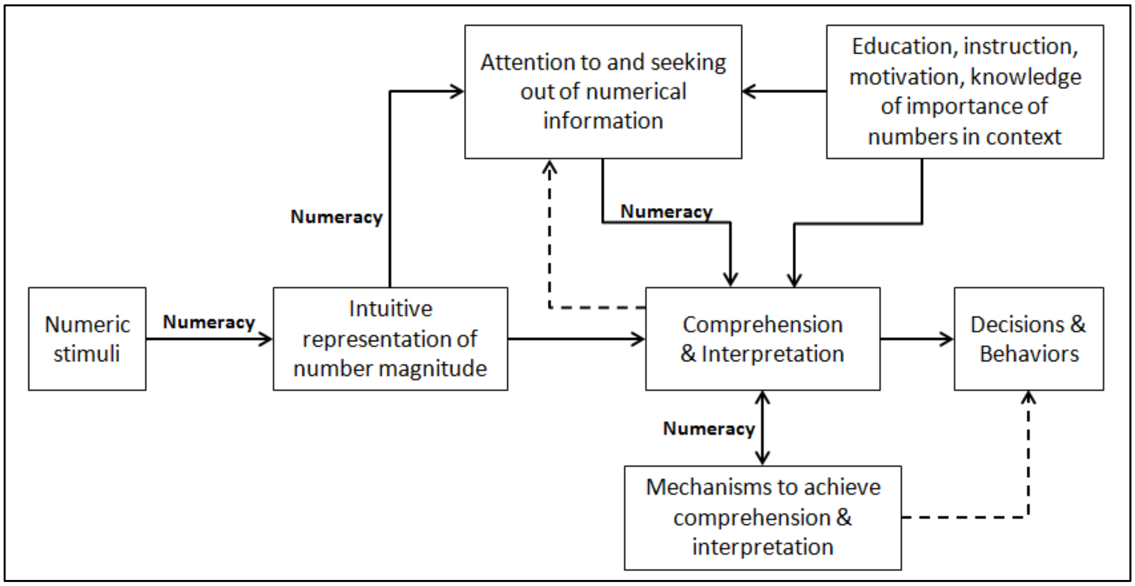

Figure 1.1. Theoretical framework of numeracy in health decision-making (Lipkus & Peters, 2009, p. 1073).

As described in the framework proposed by Lipkus and Peters (2009), several factors influence comprehension of numerical information, including the format of the numerical stimuli, the mental representation of the numerical information, and a person's level of numeracy. The format of the numerical stimuli includes quantitative formats such as percentages (e.g., 10%) and frequencies (e.g., 1 in 10) and qualitative formats such as verbal descriptions or words (e.g., unlikely) to describe the probability of an event. The mental representation of numerical information can influence comprehension because the precision of people's mental representations of quantitative values varies (Halberda, Mazzocco, & Feigenson, 2008; Peters, Slovic, Vastfjall, & Mertz, 2008). For example, a difference of 10% between two risks might be very clear to those with precise mental representations, whereas the same difference might not be so clear for those with vague or less precise mental representations of numerical information.

The role of numeracy, as suggested by Lipkus and Peters (2009), is critical in influencing comprehension and therefore, decisions and behaviors. People high in numeracy are likely to have high performance irrespective of numerical format; their mental representation of numbers is precise (Peters, Slovic, Vastfjall, & Mertz, 2008). People low in numeracy have a less precise mental representation of numbers, and comprehension of probabilities can be influenced by format (e.g., Dieckmann, Slovic, & Peters, 2009; Peters et al., 2006). For example, Peters and colleagues (2006) found that lower numerate people rated a risk presented as a percentage significantly lower than the equivalent risk presented in a frequency format.

The role of comprehension of numerical information is central to decision making and in health risk communication, but most research that has investigated the relationship between numeracy and behavior has not explicitly assessed comprehension of quantitative information. The factors that have been identified to influence comprehension include the format of the numerical stimuli, the mental representation of the numerical information, and a person’s level of numeracy. Additionally, the role of age-related differences in numeracy (e.g., Galesic, Garcia-Retamero, & Gigerenzer, 2009; Kutner, Greenberg, & Baer, 2005; Schwartz, Woloshin, Black, & Welch, 1997) has not been systematically investigated: Are there other age-related differences in cognitive abilities that influence numeracy? The focus of this dissertation was to understand how younger and older adults mentally represent and comprehend probabilities.

Format

Probabilities can be expressed in either quantitative or qualitative formats. Quantitative formats of probabilities include frequencies (e.g., 20 out of 200) and percentages (e.g., 10%). An advantage of presenting probabilities in a quantitative format is that a precise value for the probability of an event occurring is provided. A disadvantage of this format is that some minimum level of numeracy is required to understand the format. "Quantitative information is only meaningful to the extent that patients have some facility with basic probability and numerical concepts, a construct called numeracy" (Schwartz, Woloshin, Black, & Welch, 1997, p. 966).

Qualitative expressions or words can also be used to express probabilities. For example, "It is *unlikely* that a person will win the lottery." The advantage of using a qualitative format to present probabilities is that people might be more familiar and comfortable with ordinary language rather than numbers (Schwartz, Woloshin, Black, & Welch, 1997), and there is minimal to no numeracy requirement to understand qualitative expressions. A disadvantage of presenting qualitative probability expressions is that they provide less precise information than do quantitative expressions of probabilities and lend themselves to a wide range of interpretation (Karelitz & Budescu, 2004; Mazur & Hickam, 1991; Mosteller & Youtz, 1990; Windschitl & Wells, 1996; Woloshin, Ruffin, & Gorenflo, 1994). The following question then arises: In what format should probabilities be presented to maximize comprehension?

The literature is mixed regarding which format best supports comprehension of probabilistic information. Participants overestimated verbal expressions of probability as compared to frequency or percent formats when asked to assign a numerical value to the verbal expressions (Berry, Knapp, & Raynor, 2002; Knapp, Gardner, Carrigan, Raynor & Woolf, 2009; Knapp, Raynor, & Berry, 2004). However the verbal labels used were developed to express small probabilities: One descriptor expressed probabilities greater than 10% (very common) and four descriptors described probabilities less than 10% (common, uncommon, rare, very rare; see Calman, 1996). It was unclear if participants were provided with such knowledge of the scale range before estimating the quantitative value for each expression. This would likely influence estimations, as it might be that participants simply divided the range of probabilities (0-100%) by five to match each qualitative descriptor.

Other research has suggested that frequency formats provide a transparent representation of risk probabilities thereby supporting comprehension (Gigerenzer & Edwards, 2003; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). For example, Gigerenzer, Hertwig, van den Broek, Fasolo, and Katsikopoulos (2005) found a majority of participants did not understand what a 30% chance of rain meant. They argued it was because the reference class (30% of what?) was not evident in the percent format, whereas frequency formats explicitly provide the reference class and are therefore better understood. In this case, "days," was the reference class; when the weather is like today, it will rain in 3 of 10 cases.

Still other research has indicated that percent and frequency formats support comprehension similarly. Percent led to higher performance than frequency or verbal formats of probabilities in a decision making task (Dieckmann, Slovic, & Peters, 2009). An advantage of percent (e.g., 2%) and frequency (e.g., 2 in 100) over a 1 in "n" format (e.g., 1 in 50) was identified for basic mathematical operations such as comparing risk probabilities (Cuite, Weinstein, Emmons, & Colditz, 2008). It is likely that a 1 in "n" format requires higher cognitive load as the format does not follow conventional number ordering: The larger the "n," the smaller the value.

Using a qualitative approach to investigate which probability format was optimal, Schapira, Nattinger, and McHorney (2001) conducted a focus group study in which women were asked to describe their feelings about and preferences for frequencies or percentages. Many participants endorsed that frequencies were easy to understand and percentages were mathematical. However, there was evidence that not all participants understood the concept of probability given either format. Some participants were concerned with the reliability of frequencies when a low denominator was presented. This suggested that participants did not clearly understand that probabilities represent the population of people who belong to the reference class (e.g., those taking a certain medication) and that the ratio can be reduced to the lowest common denominator. This theme contradicts Gigerenzer and colleagues who advocated the use of frequencies. Other themes of confusion arose as well: One participant asked if she was supposed to identify with the 1 in 10 people or with the remaining 9 in 10; another asked how to interpret a risk of 10% as "10% of what?" (p. 462).

In the health domain, many providers include a qualitative description of risk only or in combination with quantitative probabilities (Gramling, Irvin, Nash, Sciamanna, & Culpepper, 2004; Henneman, Marteau, & Timmermans, 2008). One study found that 32% of patients wanted only numerical expressions of probability; 35.5% wanted only verbal expressions; 21.8% wanted either numbers or words, and 8.3% wanted both numbers and words (Mazur & Hickam, 1991). Other research has indicated that people prefer receiving quantitative expressions of probability rather than qualitative expressions of probability (e.g., Wallsten, Budescu, Rapoport, Zwick, & Forsyth, 1986), especially if they are higher in numeracy (Couper & Singer, 2009). However, comprehension of probabilities was not assessed in these studies, and preference does not always predict performance. To date, no study has assessed and compared comprehension of quantitative versus qualitative expressions of probabilities as a function of numeracy.

Numeracy

Low numeracy is prevalent; it impacts people of all ages across a range of education levels (Lipkus, Samsa, & Rimer, 2001; Schwartz, Woloshin, Black, & Welch, 1997; Sheridan & Pignone, 2002). In a national survey, across a broad range of domains, tasks, and number formats (whole numbers, fractions, decimals), only 13% of the 19,000 participants (less than 2,500 people) were deemed proficient in numeracy (Kutner, Greenberg, & Baer, 2005). Proficient was defined as the ability to manage and perform complex quantitative tasks that included ratio concepts and inferring operations, similar to tasks required in the health domain to obtain informed consent.

Additionally, lower numeracy has been associated with specific populations, such as older adults, the poor, and minorities (Hispanics and African Americans; Galesic, Garcia-Retamero, & Gigerenzer, 2009; Ginde, Clark, Goldstein, & Camargo, 2008; Reyna & Brainerd, 2007). Thus, a large proportion of the population is at risk for not being able to use quantitative information. The implications range from the inconvenient—not knowing when to bring an umbrella, to the life-threatening—not understanding how to manage a complex medication regimen (e.g., Apter et al., 2006; Cavanaugh et al., 2008; Estrada, Martin-Hryniewicz, Peek, Collins, & Byrd, 2004; Waldrop-Valverde et al., 2009). Thus, it is very important to determine what format leads to the best comprehension for low and high numerate people.

Mental Representation

Numeracy has also been implicated in the mental representation of numerical concepts: Peters, Slovic, Vastfjall, and Mertz (2008) suggested that higher numerate people have a more precise mental representation of quantitative values. People represent numbers spatially on a mental number line (Dehaene, 1997; Gallistel & Gelman, 2000; Moyer & Landauer, 1967) and latency measures can be used to investigate the nature of the number representations.

The greater the distance (or difference) between numbers, the easier it is for people to compare values as indicated by a decrease in response time. This is called the distance effect (Dehaene, Bossini, & Giraux, 1993). It takes more time to determine which value is greater when comparing two numbers that are very close in value (e.g., 5 compared to 6), than when two numbers are very far in value (e.g., 5 compared to 9). This distance effect suggests there is some “fuzziness” around numerical values. The idea is that the more precise a person’s mental representation of numerical values, the smaller the distance effect will be.

Peters, Slovic, Vastfjall, and Mertz (2008) modified the standard magnitude comparison task (which only included whole numbers 1 through 9) to include probabilities expressed as percentages and frequencies. The results were consistent with their hypothesis that more numerate people did in fact have a more precise representation of numerical stimuli as evidenced by a small response time slope or difference for “near” and “far” value comparisons.

Although numeracy was treated as a continuous variable and both younger and older adults were included, the data were collapsed across probability format (Peters et al., 2008). It was not clear if or how response times differed although there was reportedly no interaction of age with format. Moreover, qualitative expressions of probabilities were not investigated. An open question is if verbal expressions of probabilities are also spatially mapped to a mental number line. More research is needed to understand the relationships between numeracy and format on precision of mental representations for older and younger adults, and the relationship to overall comprehension of probabilities.

Age

The issue of determining optimal probability format to facilitate comprehension is further complicated when age-related changes in cognition are considered. Verbal knowledge increases, whereas declines in memory span, visuospatial abilities, and speed of processing have been well-documented (e.g., Park et al., 2002). Numeracy has also been negatively associated with age (e.g., Galesic, Garcia-Retamero, & Gigerenzer, 2009; Galesic, Gigerenzer, & Straubinger, 2009). Such changes in cognition might significantly influence mental representation and comprehension of probabilities across various formats. Because verbal ability is stable or even improves with age, it might be optimal to present probabilities using verbal expression to maximize comprehension for older adults.

From the numerical cognition literature, Geary and Lin (1998) did not identify age-related differences in a magnitude comparison task when comparing numbers with a magnitude greater than three (greater than the subitizing range). Geary and Lin suggested that older adults' precision of mental representations was comparable to that of younger adults. However, they indicated their results might have been due to a cohort effect in arithmetic (Schaie, 1996), thus the possibility that there are age-related declines in precision of mental representations cannot be rejected. It very well might be that older adults' educational experience created a "buffer" such that age-related declines made them appear like younger adults. Cross-sectional studies have suggested little decline in numeric ability (tests of basic mathematical skill in addition, subtraction, and multiplication); yet substantial declines have been identified in longitudinal studies (Schaie, 2007).

Additionally, age has been associated with gist rather than verbatim extraction in the reading comprehension literature (for reviews see Johnson, 2003; Meyer & Pollard, 2006). Such a difference in information processing might have an impact on comprehension of probabilities. If probabilities are represented at a gist-level, what might that look like and what would the impact on comprehension be?

Assessing Comprehension of Probabilities

Although much of the risk communication literature has focused on decision making and behavioral outcomes, some research has assessed comprehension of probabilities. Similar to the reading comprehension literature in which comprehension can be measured in various ways (Durso, Rawson, & Girotto, 2007), comprehension of risk probabilities has been operationalized in many ways including rating risk likelihood, answering questions, explaining the meaning of probability information, and recalling probabilities.

Rating risk likelihood on a Likert scale measures how a person interprets and extracts meaning from a probability. Patterns of ratings for probabilities both within and between participants can be examined to investigate relationships between ratings and formats. In a study that assessed comprehension by having university students rate the probabilities of risk on a Likert scale, risks were rated as more likely to occur when a large numerator and denominator (1,286 out of 10,000) were presented than when a small numerator and denominator were used (24.14 out of 100; Yamagishi, 1997). Participants rated a risk that had a 12.86% chance of occurring as riskier than 24.14%. These results illustrated a phenomenon called ratio bias or denominator neglect (e.g., Denes-Raj, Epstein, & Cole, 1995; Reyna & Brainerd, 2008) in which people attend only to the numerator value. This study also indicated that people do not mentally translate and represent the real number value (in this case, probabilities of 0.1286 and 0.2414) when presented with a ratio. Bonato and colleagues found a similar pattern of results with their fraction comparison tasks (Bonato, Fabbri, Umilta, & Zorzi, 2007).

This research suggests that presenting a probability in a frequency format would not be optimal for comprehension because people do not process frequencies holistically; however, percentages were not included in Yamagishi's 1997 study. In a study that did compare comprehension of probabilities by format but only included older adults, Fuller, Dudley, and Blacktop (2001) found that older adults understood percentages better than frequencies as assessed by a comprehension question. Participants were asked to indicate on a ten by ten array of figures how many people would be affected by a 20% (or 1 in 5) chance. However, the frequency format required additional processing of matching the 1 in 5 frequency to the 100 denominator, whereas the percentage denominator is always out of 100.

Comprehension questions provide a measure of the information that a person acquired from the probability and the extent to which that information is available for subsequent use (Durso, Rawson, & Girotto, 2007). Using a multiple-choice test to assess comprehension, Gigerenzer and colleagues found that people had difficulty understanding the reference class to which a single event probability belonged in the statement, “There is a 30% chance of rain tomorrow,” (Gigerenzer, Hertwig, van den Broek, Fasolo, & Katsikopoulos, 2005). The correct reference class was days: When the weather conditions are like today, in 3 out of 10 cases at least some rain fell the next day. Many participants indicated the reference class to be time (it will rain for 30% of the day) or region (it will rain in 30% of the geographic area).

Explaining probabilities can provide insight into the overall understanding and mental representation that a person has acquired; patterns of errors can be examined that might illustrate common misunderstandings. Gigerenzer and colleagues (2005) gave participants the opportunity to explain what the probability statement meant; however, there were no detailed analyses of participants’ responses to provide insight into the pattern of errors. Gigerenzer and colleagues advocated the use of “transparent” numbers such as frequencies from which people can extract the reference class easily (Gigerenzer, Gaissmaier, Kurz-Milke, Schwartz, & Woloshin, 2007); this remains an empirical question as a comparison frequency format of probability was not included in the Gigerenzer et al. 2005 study.

Recall provides a measure of the available memory representation after the comprehension process has been completed (Durso, Rawson, & Girotto). In a study that assessed comprehension of probabilities using recall (Lloyd, Hayes, Bell, & Naylor, 2001), participants (who were real patients discussing options for treatment) were informed there was between a 20% and 30% baseline risk of stroke if the surgery was not done; the surgery itself had a 2% stroke risk. Additionally, there was an 8% risk of stroke for patients three years following surgery. One month after consenting to the procedure, but before the procedure had been done, participants were asked to recall the stroke risks associated with the surgery. Two participants out of 43 indicated zero risk of stroke associated with the surgery instead of 2%, and 13 indicated they did not know the risk .

According to fuzzy trace theory (Reyna, 2008; Reyna, Nelson, Han, & Dieckmann, 2009), had participants understood the probability of the risk, they would have retained the gist or bottom-line meaning—that there was a stroke risk associated with surgery. That is, participants would have been able to describe the risks in an ordinal manner indicating which risk was more or less likely to occur than another. Fuzzy trace theory does not predict that participants will recall verbatim numerical values. Two potential confounds must be noted in the Lloyd et al. (2001) study: (1) no initial assessment of comprehension was obtained during the time of consent, and (2) the time delay of one month could have led to forgetting the risk information, which is not the same as miscomprehension.

From these studies, it is not evident how to present probabilistic expressions of risk to best support comprehension. None included qualitative expressions of probability. More evidence is needed to better understand the role of mental representations and the interaction between numeracy and format (Reyna & Brainerd, 2008). Comprehension of probabilities is difficult and more research is needed to understand the influence of format, age, and numeracy on mental representation and comprehension of probabilities.

Summary

The gaps in the research led to the high-level question: How do younger and older adults understand and mentally represent health risk probabilities? The details of how and what people understand when presented with a probability are not well understood, nor how the factors of format and numeracy influence comprehension and mental representations of probabilities for younger and older adults. The goal of the present research was to understand how these factors interact and influence comprehension of probabilities.

Dissertation Overview

A three-phase within-participant study was conducted to assess how younger and older adults mentally represented and comprehended probabilities as a function of format and numeracy. Phase 1 investigated the role of format and age by asking younger and older participants to read and discuss three health-related expository texts that each contained health risk probabilities expressed as frequencies, percents, or words. Participants' immediate and delayed descriptions of the probabilities were examined to understand how they mentally represented the probabilities. Questions about each passage were also asked to assess participants' comprehension of the probabilities. The relationship between immediate mental representation and comprehension was also assessed.

Phase 2 included the same participants and assessed mental representation and comprehension of probabilities using a magnitude comparison paradigm. Participants identified which of two probabilities presented was greater; probabilities were in the same format (frequencies, percents, or words), and participants made judgments for all formats. Precision of mental representations was operationalized as the difference in response time for near and far distance comparisons: The smaller the difference, the more precise the representation. Comprehension was assessed as accuracy on the comparison tasks by format and age.

In Phase 3, the role of numeracy on younger and older adults' comprehension and mental representation of probabilities was investigated for Phase 1 and Phase 2 using correlation and regression analyses. The goal of Phase 3 was to examine the relationships between format, age, numeracy, comprehension, and mental representations. The results of the first two phases were combined to provide an overall picture of how probabilities were mentally represented and comprehended as a function of format, age, and numeracy.

CHAPTER 2: PHASE 1 OVERVIEW

Research Questions

The aim of Phase 1 was to investigate the following questions:

- How do younger and older adults mentally represent probabilities and do those representations vary as a function of format and age?
- How does format influence comprehension of health risk probabilities for younger and older adults? Is the same format ideal for both younger and older adults?
- What was the relationship of comprehension of probabilities with mental representation, format, and age?

Method Overview

A multiple-measure approach was taken to investigate how people mentally represented and comprehended health risk probabilities as a function of format and age. Participants read and discussed fictional health-related expository passages containing health risk probabilities. A teach-back approach combined with comprehension questions, and tests of recall and recognition were used to identify how mental representation and comprehension of health risk probabilities was influenced by format and age.

Teaching-back information provides insight into the comprehension of probabilities as the person must read the information, prepare to explain it by integrating the information, and then actually explain it (Doak, Doak, & Root, 1996; Schillinger et al., 2003; Quickguide to Health Literacy, accessed 2010). This approach also provides insight into how participants mentally represent probabilities, as their teach-back discussion can be examined for patterns within and between participants.

Healthcare providers have advocated the teach-back approach as a valuable method to assess patient comprehension (Fink et al., 2010). For example, a physician might say to a patient, “To make sure that I was being clear enough, could you please share with me the main points you got from our discussion?” (Farrell et al., 2009, p. 129). Patients should be able to demonstrate their understanding of the information by explaining what they have just read or heard; a verbatim repetition does not necessarily indicate comprehension (Durso, Rawson, & Girotto, 2007).

Another measure of comprehension used was accuracy on multiple-choice questions about health risk probabilities; this provided insight into the information that participants acquired from the passage (Durso et al.). Explicit and inferential questions were used to assess comprehension (Morrow et al., 2005).

Additionally, recall tests assessed the memory trace of the probabilities that was still available after reading, discussing, and answering questions (Durso et al.). The delayed cued recall test explicitly asked participants to provide the risk probability expressions for each problem (mild, moderate, and severe) for each health passage. Lastly, the delayed recognition test was used to examine if the participants understood that mild problems would be experienced by the most people for all health treatments and severe problems by the fewest people. This was a very high-level assessment of gist (Reyna, 2008).

Hypotheses

Immediate Mental Representation of Probabilities

Age differences in accuracy of immediate mental representations of probabilities might emerge, as older adults tend to extract the gist or bottom-line meaning from passages, whereas younger adults retain the surface or verbatim level of information (e.g., Johnson, 2003; Meyer & Pollard, 2006). Younger adults' accuracy was expected to be high. It might be that older adults use less specific terms (gist-level) to describe the high level idea of the probability of a problem occurring at a binary level: The problem might occur. The data might indicate what gist extraction of quantitative and qualitative probabilities would be. Additionally, because older adults maintain high verbal knowledge, it was expected that the high accuracy would be achieved in the words format.

Comprehension of Probabilities

It was expected that a difference in accuracy of comprehension would emerge by format. Gigerenzer and colleagues' research would be supported if participants were most accurate with frequencies; the Fuller et al. (2001) findings would be supported if participants were most accurate with percents. Because verbal ability is stable or even improves with age, it might be optimal to present probabilities using words (qualitative verbal descriptors) to maximize comprehension for older adults. If words were found to support highest accuracy, it would be a novel empirical finding: There is no evidence to suggest expressing probabilities as words best supports comprehension.

Relationship between Immediate Mental Representation and Comprehension

Immediate mental representations were also examined for their relationship to performance on comprehension questions. A strong positive relationship between the accuracy of the immediate mental representation of the probability and accuracy on the comprehension questions was expected.

Delayed Mental Representation of Probabilities

It was hypothesized that participants would correctly recall the ordinal ranking of problems (i.e., mild > moderate > severe problems) but not the verbatim probabilities across the three passages. This would be consistent with fuzzy trace theory (Lloyd, Hayes, Bell & Naylor, 2001; Reyna, 2008; Reyna, Nelson, Han, & Dieckmann, 2009).

As participants read three passages each with three expressions of probability in three different but equivalent formats, it was expected that participants' recall would be accurate. That is, if participants understood the probabilities, each probability would be reinforced with each presentation despite the format change. Poor cued recall performance would suggest that participants were unable to translate between equivalent formats and that they did not understand the probabilities. Age-related differences were not expected because of the multiple presentations of different but equivalent probabilities.

CHAPTER 3: PHASE 1 METHOD

Participants

Thirty-nine younger adults between the ages of 18 and 28 years were recruited from the Georgia Institute of Technology undergraduate pool and compensated with course credit. Three participants were excluded and replaced (One was ill; two due to experimenter error). The data reported are for 36 younger adult participants. The mean age for the younger adult group was 20 years (*SD*=2.2) Twenty participants were female; sixteen were male.

Forty-two older adults between the ages of 65 and 75 years were recruited from the Human Factors and Aging Laboratory database and were compensated $50 for their time. The data from six older adult participants were excluded and replaced (3 were outliers; 1 incomplete; 2 due to experimenter error). The data reported are for 36 older adult participants. The mean age for the older adult group was 71.1 years (SD=2.4). Twenty-three participants were female; thirteen were male. Eighty-five percent of older adults (*n*=31) had some college or higher education.

Table 3.1 describes younger and older participants' racial groups.

Table 3.1

Percent of Participants by Racial Group

	Racial Group						
	White Caucasian	Black/ African American	Asian	Native Hawaiian/ Pacific Islander	Multi-Racial	No primary group	Total
Younger Adults	64%	8%	22%	3%	3%	0%	100%
Older Adults	61%	31%	0%	0%	3%	6%	101%*

*Does not sum to 100% due to rounding error.

Materials and Procedure

Figure 3.1 provides an overview of the order and timing of the specific tasks performed in Phase 1. The details of each task are provided next.

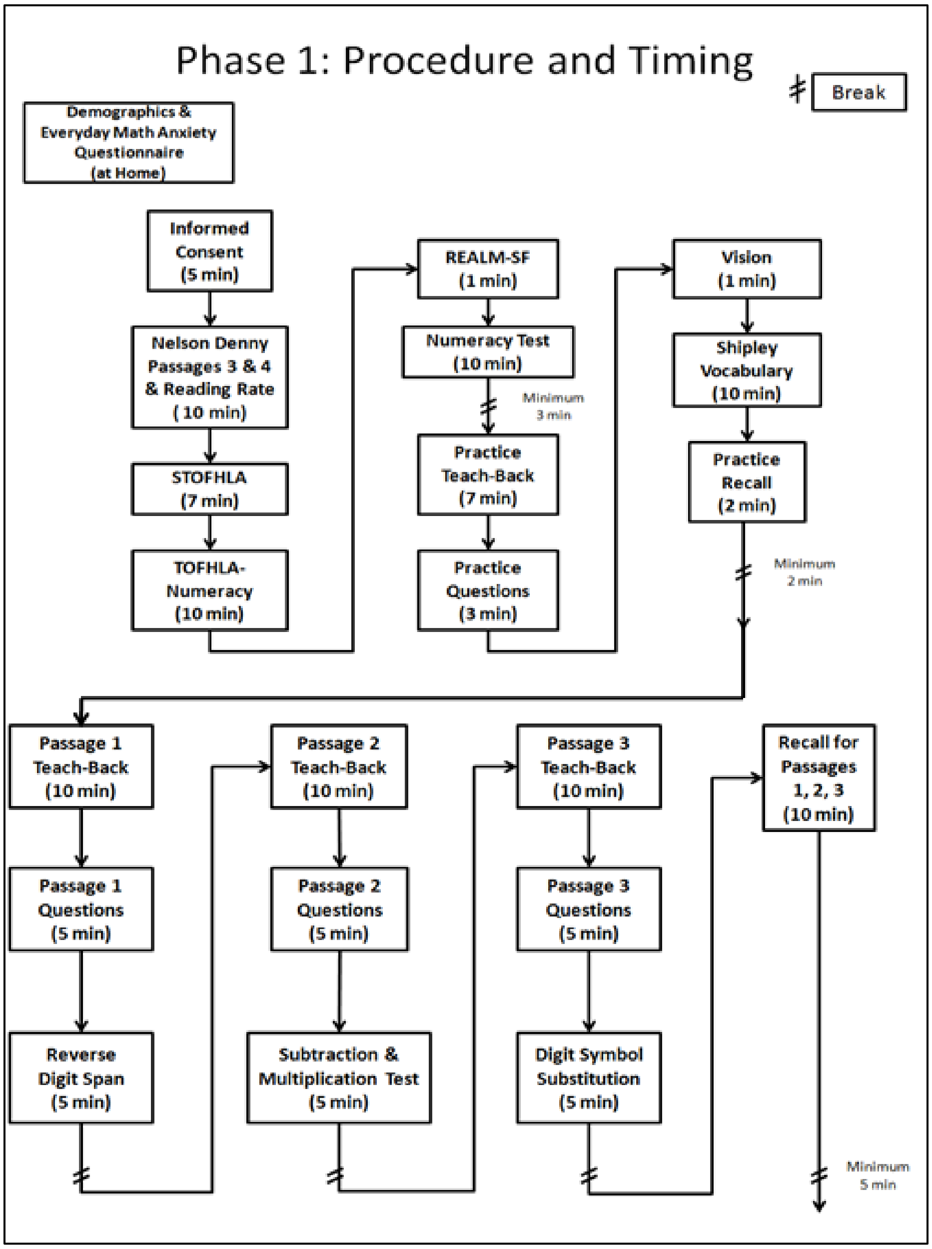

Figure 3.1. Procedure and timing details for Phase 1.

Questionnaires and Ability Tests

Demographic and general health information were collected for each participant, and all completed an Everyday Math Anxiety Questionnaire (locally developed; see Appendix A). The ability tests are listed and described in Table 3.2.

Table 3.2

Ability Tests

Ability Test		Max Score	Reference
Nelson Denny Reading Test[a]	Reading Comprehension	10	Brown, Fischco, & Hanna (1993)
Nelson Denny Reading Test	Reading Rate	610	Brown, Fischco, & Hanna (1993)
S-TOFHLA[b]	Health Literacy	36	Nurss, Parker, Williams, & Baker (2001)
TOFHLA-Numeracy	Health Numeracy (Basic)	17	Nurss, Parker, Williams, & Baker (2001)
REALM-SF	Ability to read medical terms	7	Arozullah et al. (2007)
Numeracy	General and health probability	12	Lipkus, Samsa, & Rimer (2001); Schwartz, Woloshin, Black, & Welch (1997)
Shipley Institute of Living Scale	Semantic Knowledge (Vocabulary)	40	Shipley (1986)
Reverse Digit Span	Memory Span	14	Wechsler (1997)
Subtraction & Multiplication Test N3[c]	Math Ability (Basic)	60	Ekstrom, French, Harman, & Derman (1976)
Digit-Symbol Substitution	Perceptual Speed	100	Wechsler (1997)

[a] Only passages 3 and 4 used.
[b] Prompts modified to reflect current dates to minimize distraction (e.g., 1993 changed to 2011").
[c] Half of practice problems deleted to match timing to the reverse digit span and the digit symbol substitution tasks.

Health Passages and Health Risk Probabilities

Health passages. Three fictional health-related expository passages were created for this study. The passages provided information about and explained risks associated with the following: Medication G, Procedure A, and Vaccine Q. Passages were presented in 14-point bold black sans serif font type (Calibri) on white paper. Table 3.3 describes passage source.

Table 3.3

Health Passage Sources

Passage	Source
Medication G	Motrin: http://www.drugs.com/motrin.html Motrin PM: http://www.drugs.com/sfx/motrin-pm-side-effects.html
Procedure A	Appendectomy Informed Consent Form: http://www.dialogmedical.com/content/sample-documents/
Vaccine Q	CDC Tdap Vaccine Information Statement: http://www.cdc.gov/vaccines/pubs/vis/default.htm#tdtdap

Each passage comprised seven sections. The first three sections described the treatment; the fourth section described a general risk; the fifth, sixth, and seventh sections described mild, moderate, and severe problems, respectively. Passages were equated by Flesch-Kincaid reading level and number of words per section. Overall, the average Flesch-Kincaid reading grade level of the passages was 8.7. The average word count was 225 (range: 217-234). See Appendix B for details.

Health risk probabilities. The probability of experiencing mild, moderate, and severe problems was expressed as frequency, percent, or words. These risk expressions were in the same format for each passage; the format varied between passages. One passage expressed the risks of mild, moderate, and severe problems using frequency, another passage used percent, and another passage used words. The passages were fully counterbalanced for format (percent, frequencies, words) and passage topic (medication, procedure, vaccine). See Appendix C for the counterbalancing scheme. The values of the probabilities for mild, moderate, and severe problems can be found in Table 3.4. The word expressions used were taken from Windschitl and Wells (1996; refer to Appendix D Figure 1 for the complete list of verbal expressions and their numerical equivalents).

Table 3.4

Probability Values by Format and Problem

Problems	Frequency	Percent	Words
Mild	17 in 20	85%	very likely
Moderate	1 in 5	20%	quite unlikely
Severe	1 in 10,000	0.01%	almost totally impossible

The probability values were modified from those contained in the CDC Tetanus, Diphtheria (Td) or Tetanus, Diphtheria, Pertussis (Tdap) vaccine information statement. Table 3.5 provides an example of a health passage. According to the CDC statement, mild, moderate, and severe problems could occur in up to 80% (8 in 10), 6.25% (1 in 16), and 0.0001% (1 in 1,000,000) of people who got the vaccine, respectively. These values were modified for the following reasons:

1. Using Windschitl and Wells (1996) values for verbal labels of probabilities, there were not separate labels for 6.25% and 0.0001%: Both values would have been "almost totally impossible." The quantitative values assigned to verbal expressions of probability were also consistent with those described in Mazur and Hickam (1991) and Mosteller and Youtz (1990).
2. The mild and moderate values were selected to be approximately the same value away from 50% (mild was +35%; moderate was -30%).
3. The word labels for mild and moderate problems should contain both a modifier and a "likely" root (one with "likely" and the other with "unlikely"); the modifier should not be the same.
4. All frequency values were non-reducible.

5. Frequency values were selected to have unique denominators preventing participants from comparing the numerators only.

Table 3.5

Example Health Passage (Vaccine Q Information Statement with Percent Format)

Vaccine Q Information Statement

Why get vaccinated?

Infants and children are routinely vaccinated against Disease Q. But older children, adolescents, and adults need protection from this disease too. Vaccine Q provides that protection.

The United States averaged more than 100,000 cases of Disease Q each year before the vaccine. Since the vaccine has been available, Disease Q cases have fallen significantly.

Disease Q causes pain, a rash, a high fever, and it can be deadly. Disease Q is spread from person to person. Vaccine Q strengthens the body's ability to fight off Disease Q.

What are the risks from Vaccine Q?

With Vaccine Q, as with any medicine, there is always a risk of an allergic reaction. However, getting Disease Q would be much more likely to lead to severe problems than getting the vaccine.

<u>85%</u> of adults will experience mild problems after getting Vaccine Q. These problems are noticeable but do not interfere with activities. An example of a mild problem is a low-grade fever.

<u>20%</u> of adults will experience moderate problems after getting Vaccine Q. These problems interfere with activities but do not require medical attention. An example of a moderate problem is pain at the injection site.

<u>0.01%</u> of adults will experience severe problems after getting Vaccine Q. These problems require medical attention. An example of a severe problem is seizures.

Note: Underlined sections varied according to format condition (i.e., frequency, percent, words). See Appendices E-M for complete set of stimuli.

<u>Teach-Back Instructions</u>

Participants were told that that they were going to read and evaluate three fictional health-related passages for clarity. Participants listened to each passage section as the experimenter read it aloud; they were then asked to read each section again to themselves before performing the teach-back. The instructions were structured to minimize listener effects as research has suggested that the identity of the listener can greatly influence how people re-tell stories or information (e.g., Adams, Smith, Pasupathi, & Vitolo, 2002; Hyman, 1994). Each participant was instructed as follows,

> I will give you brief sections of the passage to teach-back or explain one section at a time. Remember, the idea is to imagine that you have to explain this information to a friend of similar age and background as yourself. Please describe the critical pieces of information that you understood from reading each section. Take as much time as you need. When you are ready to teach-back or describe the critical pieces of information please turn the paper over.

Participants were told their teach-back responses would be audio-recorded for analysis.

Practice Passage

A practice passage was presented before the experimental passages to give participants practice performing the teach-back procedure. See Appendix N for the passage and Appendix O for the questions and correct answers.

Comprehension Questions

Participants answered 20 comprehension questions about each passage (10 each about general and probability content) and were allowed to refer back to the passage as needed to answer the questions. The questions were further divided into three explicit and seven inferential questions for each content type. Answers to explicit questions were contained within the passage; whereas answers to inferential questions were not explicitly contained within the passage. For the general content inferential questions, participants had to infer from the passage what the best answer would be. For the probability content inferential questions, participants had to translate between probability formats. Questions were randomized with the rule that no more than three of the same content questions could be presented in a row. Refer to Appendices P-R for the comprehension questions for each health passage and Appendix S for the answers.

Delayed Free Recall

Once all passages were taught back, participants were asked to describe what they could remember from each passage in as close to the same words from the passage as they could. Participants were prompted with the topic of each passage in the same order in which they were read and taught-back.

Delayed Cued Recall

Participants were then asked how often each of the problems (mild, moderate, and severe) was experienced according to each of the health passages.

Delayed Recognition

Participants then identified which problems (mild, moderate, or severe) affected the most people and the fewest people according to each passage.

A minimum of a five-minute break was given to the participants before Phase 2 started. All participants completed Phase 1 and Phase 2 on the same day except for one older adult who returned to the lab on another day to finish Phase 2. This participant had recently fallen and had a sore leg; sitting was becoming painful by the end of Phase 1.

Design

The experiment was a 3 (Format: frequency, percent, words) x 2 (Age: younger adults, older adults) quasi-experimental split plot design. The format variable was manipulated within-subjects; age served as a grouping variable. The dependent variable was accuracy on comprehension questions. Participants' teach-back responses and delayed cued recall responses were analyzed for patterns of errors by format and age.

CHAPTER 4: PHASE 1 RESULTS

Analysis Overview

Format was manipulated within participant (frequency, percent, words) and age served as a grouping variable (younger, older). Alpha was set at .05. The primary dependent variable was accuracy on comprehension questions. A mixed model analysis of variance conducted with format (frequency, percent, words) entered as a within participant variable and age as a grouping variable.

Immediate Mental Representation of Probabilities

Acknowledgment of Probabilities

Teach-back responses were analyzed to understand how participants mentally represented probabilities. First, participants' inclusion of probabilities for mild, moderate, and severe problems in their teach-back responses was examined to understand if the probabilities were mentally represented at all. Acknowledgment of probabilities was defined as any indication of the problem occurring. Expected values for acknowledgement of probability for each age group was as follows: 36 participants x 3 probabilities (for mild, moderate, severe problems) = 108. Table 4.1 shows that younger and older adults acknowledged probabilities for mild, moderate, and severe problems, irrespective of the format in which the probabilistic information was given. These data suggest that participants attended to the probabilities and considered them as critical pieces of information: Participants were instructed to explain the critical pieces of information they understood from the passage.

Table 4.1

Acknowledgment of Probabilities by Age Group and Format

	Younger Adults			Older Adults		
Acknowledgment	Frequency	Percent	Words	Frequency	Percent	Words
No	0%	0%	0%	4%	5%	6%
Yes	100%	100%	100%	96%	95%	94%

Accuracy of Immediate Mental Representations

Participants' teach-back responses were examined to determine how accurately they mentally represented each probability. A lenient scoring rubric was used to examine the results because participants were instructed to explain the critical pieces of information they understood. They were not instructed to use the same words as in the passage; it was more important to determine whether participants were mentally representing the probabilities "within the ball park" of the probability value given. (Note: Data patterns were similar for strict and lenient scoring; however, participants had higher accuracy scores in the lenient scoring analysis.)

Scoring Rubric for Accuracy

Correct responses had to be within a range of the given value. Table 4.2 shows the scoring rubric. The values of verbal expressions were extrapolated from the literature (Mazur & Hickam, 1991; Mosteller & Youtz, 1990; Windschitl & Wells, 1996).

Table 4.2

Scoring Rubric for Accuracy of Immediate Mental Representations of Probabilities

	Mild Problems (85%; 17 in 20; very likely)	Moderate Problems (20%; 1 in 5; quite unlikely)	Severe Problems (0.01%; 1 in 10,000; almost totally impossible)
Acceptable range	Greater than 50% but less than 100%	Greater than 1% but less than 50%	Less than 1%
Other acceptable responses	most people many people a lot of people majority common high probability pretty likely large percentage high number	some people minority not likely unlikely	nearly impossible almost impossible few people rare low likelihood small percentage seldom

Proportion Correct by Format and Age Group

Figure 4.1 shows the proportion correct of teach-back representations by format and age group. A mixed model analysis of variance revealed a main effect for format ($F(2, 69) = 15.0$, p <.001, $\eta_p^2 = .30$). Follow-up paired comparisons identified that accuracy for the frequency and percent were significantly higher than for the words format (p < .001 for both). The data indicate that format influenced accuracy of representations; accuracy was lowest for the words format for both younger and older adults. Younger adults were more accurate than older adults across all formats ($F(1, 70) = 23.2$, p <.001, $\eta_p^2 = .25$).

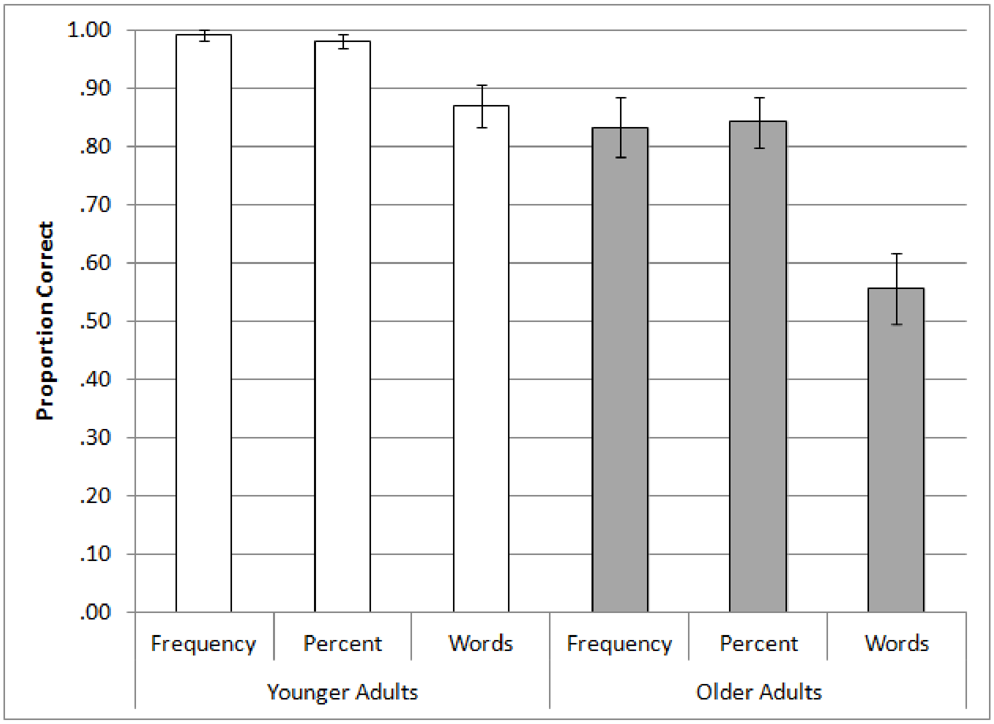

Figure 4.1. Proportion correct by format and age group for immediate representations of probabilities. Error bars represent standard error of the mean.

The age difference was expected; younger adults tend to maintain verbatim or surface representations of text (e.g., Meyer & Pollard, 2006). However, the effect of format on accuracy was not expected. It was hypothesized that older adults would do as well if not better with the words format; the data indicated otherwise. To investigate why the representations were most inaccurate for the words format, the data were examined at the probability level. Differences between younger and older adults' responses were also examined.

<u>Proportion Correct by Format, Age Group, and Probability</u>

The data were examined to understand if there were particular probabilities that were driving the inaccurate mental representations by format and age group. Table 4.3 shows the percent correct by format, age group, and probability.

Table 4.3

Percent Correct by Age Group, Format, and Probability

	Frequency			Percent			Words		
	17 in 20	1 in 5	1 in 10,000	85%	20%	0.01%	very likely	quite unlikely	almost totally impossible
YA	100%	100%	97%	100%	100%	94%	81%	92%	89%
OA	77%	89%	94%	91%	97%	76%	47%	59%	73%

YA=Younger Adults
OA=Older Adults

Younger adults were highly accurate representing both frequency and percent formats; they were less accurate in representing words format probabilities. In particular, younger adults did not correctly represent the verbal probability “very likely;” there was a low percent correct compared to their performance on other formats. To understand why accuracy was lower for “very likely,” the errors were reviewed to examine if any patterns emerged.

Six younger adults expressed the probability “very likely” with a binary term (e.g., “might,” “may,” or “possible”). These expressions suggest that these participants were translating “very likely” into something akin to a binary response (i.e., mild problems may or may not happen). They were acknowledging that the problem could occur but without any indication of likelihood. One participant stated, “There are mild side effects,” to represent “very likely.” Such a statement does not indicate a probability; rather it incorrectly indicates that every person will experience mild problems. This pattern suggests that these younger adults were less exact in representing the verbal phrase, “very likely.”

Table 4.3 shows that older adults had low accuracy for the following probabilities: "17 in 20," "0.01%," "very likely," "quite unlikely," and "almost totally impossible." Older adults incorrectly represented "17 in 20" most often for the frequency probabilities. This might be because "17 in 20" is not a commonly used fraction, whereas "1 in 5" and "1 in 10,000" use denominators of 5 and a base of 10 which are common in our culture (e.g., Burkell, 2004). The incorrect expressions were examined to understand participants' mistakes.

Six older adults said, "17 **to** 20" which might have been a mistake substituting "to" for "in," but the fact that six older adults said this suggests it is more a pattern of representation errors rather than a slip of the tongue. The expression "17 to 20" suggests that 17 to 20 people will experience mild problems rather than using a probabilistic expression that discusses chance or likelihood based on a population.

One older adult said, "some," which suggests that the person understood that 17 in 20 did represent more than zero, but the expression "some" (according to the scoring rubric) was too inexact—17 in 20 represents "most" people not "some" people. Similar to the younger adults, one older adult said, "There are mild symptoms," which does not indicate chance; rather it incorrectly suggests that every person will experience mild problems.

With respect to the percent format, older adults had the lowest accuracy for "0.01%." A decimal in a percentage might be uncommon for some people and difficult for older adults to understand. The incorrect expressions that older adults used were examined for patterns. Five older adults said, "1%," and one said "01%." These representations are lacking the important decimal. The participant who said, “01%,” tried to represent the zeros, but did not understand exactly what the zeros meant.

Two older adults did not include an indication of percent. One older adult stated, “0 to 1” This suggests that the participant represented components of the 0.01% value, but did not represent it correctly. “0 to 1” suggests that 0 to 1 person will experience severe problems rather than a probability or proportion of all of the people who receive the treatment. The other participant said, “Less than .01 of adults.” It might be that the person represented the value as a rate rather than a percentage, but the conversion was incorrect as percent represents a divisor of 100, thus the correct rate would have been 0.0001.

Probabilities presented in the words format led to the worst performance for older adults; they did not correctly represent word probabilities in their teach-back responses. To express the probability of “very likely,” twelve participants used binary terms (e.g., “might,” “may,” or “can”), four used terms of certainty (e.g., “will have,” “there are”), and two older adults used incorrect expressions (i.e., “some,” “very unlikely”). With respect to the low accuracy for immediate representations for “quite unlikely,” seven older adults used binary terms (e.g., “may,” or “could”), four used terms of certainty (e.g., “will experience”), and three used incorrect expressions (i.e., “quite likely” and “probably will”). To express “almost totally impossible,” seven participants used incorrect expressions (e.g., “impossible,” “totally likely,” “some”), one used a binary term (“can cause”), and one used a number without a percent (“0.001”) and no other indication of rate or probability. This pattern suggests that older adults are less exact with words (i.e., using the binary terms), or that words do not clearly indicate a probability value (i.e., using incorrect expressions).

Summary of Accuracy of Immediate Mental Representations of Probabilities

The majority of participants acknowledged probabilities in their immediate mental representations. Format influenced accuracy of mental representations, such that accuracy was higher for frequencies and percents than for words for both younger and older adults. Younger adults were more accurate overall than older adults. The data did not support the hypothesis that older adults would be more accurate with words than frequencies or percents. The data did suggest that older adults described probabilities at a gist level, using terms such as, “might” and “could happen.”

Comprehension of Probabilities

Gist-Level Comprehension

Participants identified in a multiple-choice test which problem (mild, moderate or severe) would be experienced by the most and fewest people according to each passage for a total of six questions. For all passages, mild problems were experienced by the most people and severe by the fewest. Total mean scores and standard deviations were 5.9 (0.23) for younger adults and 5.0 (1.5) for older adults. Table 4.4 shows how many questions younger and older adults answered correctly. Figure 4.2 shows the mean score for each format by age group.

Table 4.4

Delayed Recognition Test: Percent Correct by Age Group

Number Correct	Younger Adults	Older Adults*
6	94%	56%
5	6%	19%
4	0	11%
3	0	3%
2	0	6%
1	0	6%

*Does not sum to 100% due to rounding error.

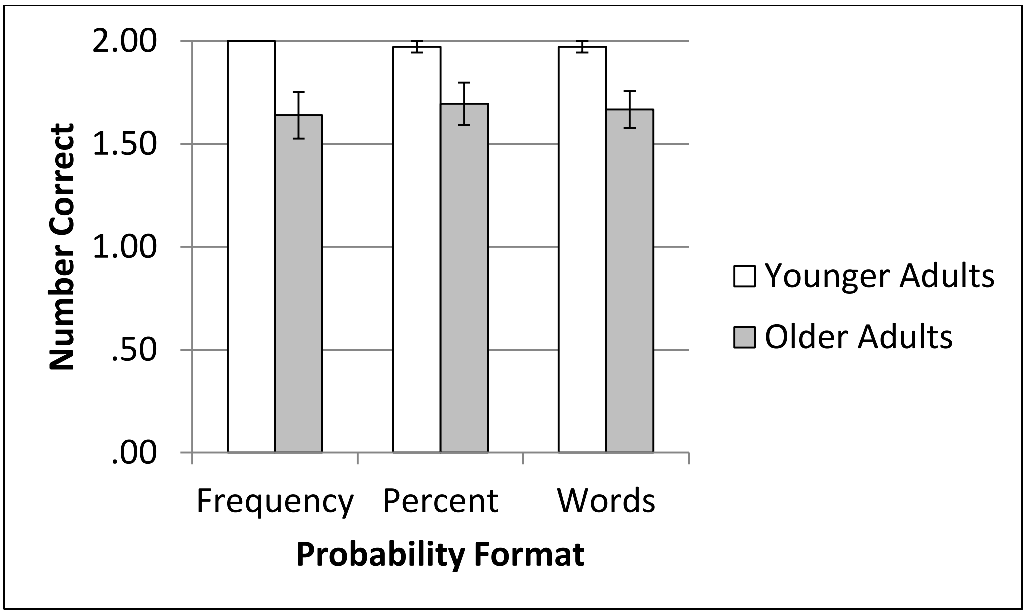

Figure 4.2. Mean score for younger and older adults for each format. Error bars represent standard error of the mean.

A mixed model analysis of variance revealed only a significant main effect for age ($F(1, 70) = 14.1$, $p < .001$, $\eta_p^2 = .17$). This measure contained only six questions and provided only a high-level assessment of gist-level comprehension. It illustrated that the majority of participants (54 out of 72 participants; 75%) understood at a gist level that mild problems would be experienced by the most people and severe by the fewest. However, the results are far from encouraging as 25% of the participants did not correctly understand the gist of the health risk probabilities (scored 5 or lower); this was driven by older adults' performance.

This measure was deceiving as it did not measure deeper comprehension, and it was confounded with memory decay. Comprehension is a multi-dimensional construct, thus more detailed assessments of comprehension are needed to understand under what conditions people do and do not successfully processing health risk probabilities. The goal of this part of the study was to investigate the role of format on comprehension and if there were age-related differences.

Probability Comprehension

Comprehension was assessed via 20 multiple-choice questions that participants answered after reading and discussing each passage. Ten questions served as filler and assessed general content comprehension. An ANOVA yielded a main effect of age ($F(1\ 70) = 28.4$, $p < .001$, $\eta_p^2 = .29$). There was no difference in general comprehension as a function of format; this serves as a type of manipulation check that probability format did not influence comprehension of the other information in the passage.

An ANOVA of the 10 probability content questions revealed a main effect for both format ($F(2, 69) = 10.3$, $p <.001$, $\eta_p^2 = .23$) and age ($F(1, 70) = 63.3$, $p < .001$, $\eta_p^2 = .48$). Follow-up paired comparisons identified that accuracy for the frequency and percent were significantly higher than for the words format ($p < .001$ and $p < .05$, respectively). See Figure 4.3.

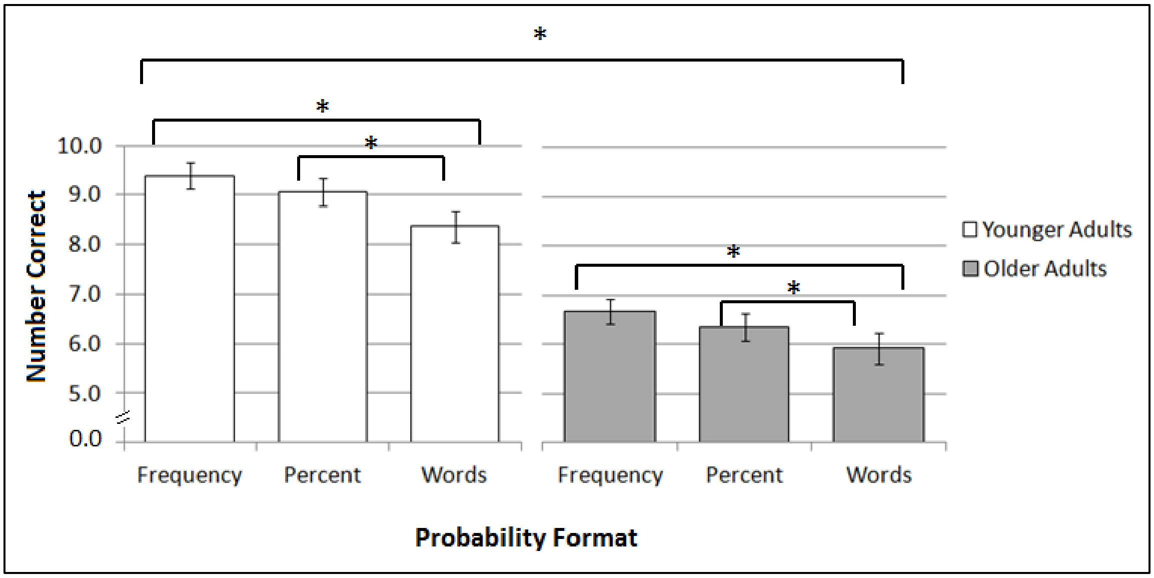

Figure 4.3. Mean number correct for comprehension questions by format for older and younger adults. Error bars represent standard error of the mean.

The comprehension questions were used to measure both explicit and inferential aspects of comprehension and provided more details about participants' comprehension than the simple delayed recognition test. The results suggest that format influences comprehension of probabilities such that words lead to lowest comprehension. To gain a deeper understanding of the problem space of comprehension, the relationship between accuracy of immediate representation and comprehension of probabilities was examined.

Relationship between Immediate Representation and Comprehension

Table 4.5 and Figure 4.4 show the mean number of correct questions by accuracy of representation and format for younger and older adults.

Table 4.5

Comprehension Question Accuracy as a Function of Immediate Representation, Format, and Age Group

	Younger Adults			Older Adults		
Representation	Frequency *M (SD)*	Percent *M (SD)*	Words *M (SD)*	Frequency *M (SD)*	Percent *M (SD)*	Words *M (SD)*
Correct*	9.4 (0.9)	9.0 (1.0)	8.7 (1.3)	7.3 (1.7)	7.0 (1.7)	7.3 (1.5)
Incorrect	10.0 (N/A)	9.5 (0.7)	7.6 (1.2)	4.9 (2.0)	5.4 (2.4)	5.4 (2.4)

*Correct = All three probabilities were represented correctly.

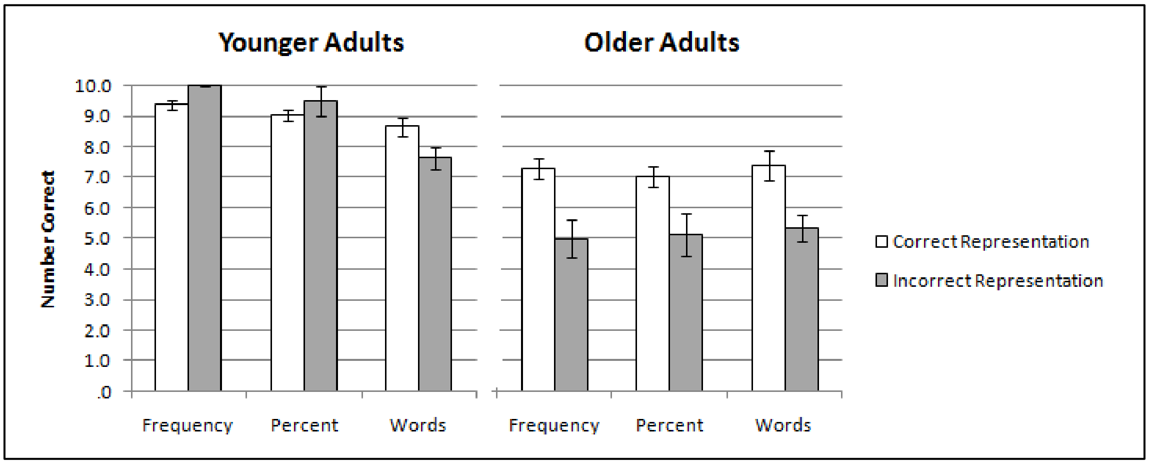

Figure 4.4. Mean number correct for comprehension questions by immediate representation for older and younger adults. Error bars represent standard error of the mean.

Point bivariate correlations were conducted for each format by age group to examine the relationship between accuracy of immediate mental representation and accuracy on comprehension questions. Table 4.6 shows the Pearson correlation value for younger and older participants for each format. Accuracy of immediate mental representation was input as a nominal variable (1 = yes; 0 = no).

Table 4.6

Correlation between Accuracy of Immediate Mental Representation and Accuracy on Comprehension Questions by Format and Age

	Frequency	Percent	Words
Younger Adults	*r*(36) = -.12	*r*(36) = -.11	*r*(36) = .36*
Older Adults	*r*(36) = . 52*	*r*(36) = .43*	*r*(36) = .41*

*p<.05

The data suggest a weak negative relationship for younger adults for the frequency and percent formats. Those few younger adults who incorrectly represented the immediate representations of frequency and percent (n = 1 and 2, respectively) had higher accuracy on the comprehension questions. Consistent with expectations, a strong significant positive correlation for words for younger adults was identified. The eleven younger adults who incorrectly represented words had lower accuracy on comprehension questions.

The significant positive correlations for older adults across all formats suggest that immediate representations are strongly related to performance on comprehension questions. This difference in relationships between younger and older adults for the frequency and percent formats was likely due to the low number of younger adults who were incorrect in their immediate representations.

Delayed Mental Representation of Probabilities

Delayed Acknowledgment of Probabilities

Delayed representations of probabilities can provide insight into participants' mental representations of the probabilities after three exposures to the same probabilities in different formats. First, participants' inclusion of probabilities for mild, moderate, and severe problems was examined to understand if the probabilities were mentally represented at all. Acknowledgment of probabilities was defined in the same manner as for immediate mental representation. Table 4.7 shows percent of participants who acknowledged probabilities in their delayed mental representations by format and age group.

Table 4.7

Delayed Representation of Probabilities by Format and Age Group

	Younger Adults			Older Adults		
Acknowledgment	Frequency	Percent	Words	Frequency	Percent	Words
Missing	0	0	3%	6%	6%	11%
No	0	0	0	8%	11%	10%
Yes	100%	100%	97%	86%	83%	79%

The data suggest that almost all of the probabilities were acknowledged by the younger adults, whereas older adult participants did not acknowledge several of the probabilities. They either answered the questions inappropriately (e.g., see your doctor) (scored as "No") or chose not to answer the question (scored as "Missing"). This might suggest that older adults had a very "weak" or no mental representation of the probabilities.

Although there are age-related differences in recall such that older adults are worse than younger adults, the participants were exposed to each probability three times in various formats. Had they understood this, it would be much more likely that they would have been able to mentally represent something. This could imply that many older adult participants did not understand the probabilities were equivalent across formats.

Accuracy of Delayed Mental Representations of Probabilities

Gist Accuracy of Delayed Mental Representations

The gist of the probabilities was investigated by examining the ordinal ranking of probabilities by format. Namely, did participants recognize that the probabilities for mild problems should be greatest and the probabilities for severe problems least? Table 4.8 shows the percent of participants who correctly rank ordered the probabilities by format and age group.

Table 4.8

Gist Accuracy of Delayed Mental Representations of Probabilities by Format and Age Group

	Younger Adults			Older Adults		
	Frequency	Percent	Words	Frequency	Percent	Words
Correct	**100%**	**100%**	**92%**	**53%**	**67%**	**58%**
Incorrect	0	0	6%	31%	14%	17%
Missing	0	0	3%	17%	19%	25%
Total	100%	100%	101%*	101%*	100%	100%

*Does not sum to 100% due to rounding error.

The data suggest that overall performance was high across all formats for younger adults. Older adults' gist accuracy was much lower than younger adults across all formats. Of the three formats, percent appears to support gist accuracy best for older adults; 67% of participants correctly ordered percent values, compared to 53% for frequency and 58% for words.

Accuracy of Delayed Mental Representations

Participants' delayed mental representations were evaluated individually for accuracy of the probability. The same lenient scoring rubric used to evaluate the immediate representations was used. Table 4.9 shows the proportion correct by format and age group of the delayed representations of the probabilities. A mixed model analysis of variance revealed a main effect for age ($F(1, 70) = 6.90$, $p < .001$, $\eta_p^2 = .39$).

Table 4.9

Proportion Correct by Age Group and Format

	Frequency *M (SD)*	Percent *M (SD)*	Words *M (SD)*
Younger Adults	.99 (.06)	.98 (.08)	.95 (.21)
Older Adults	.65 (.36)	.60 (.35)	.57 (.38)

The data suggest that format did not influence accuracy of delayed mental representations; the pattern was similar between younger and older adults. Overall accuracy for younger adults was higher than older adults across all formats. Younger adults' accuracy of delayed mental representation of words was higher than their accuracy for immediate mental representations ($M = .87$). Older adults' accuracy of delayed mental representations was much lower than for immediate mental representations for frequency ($M = .83$) and percent ($M = .84$), whereas they were comparable for the words format ($M = .56$). These data are consistent with the aging literature that show older adults have lower recall accuracy than younger adults (for a review see Hoyer & Verhaeghen, 2006).

Summary of Delayed Mental Representation of Probabilities

The majority of participants acknowledged probabilities in the delayed mental representations, though older adults were more likely to decline to answer or to give inappropriate responses. Unlike the data pattern for immediate mental representation, the delayed mental representation data suggested that format did not influence accuracy of delayed mental representations. Younger adults were high in accuracy across all formats; older adults were much worse.

CHAPTER 5: PHASE 1 DISCUSSION

The aim of Phase 1 was to examine the influence of probability format on mental representation and comprehension for younger and older adults. The teach-back method was used to elicit participants' mental representations of probabilities. These immediate mental representations also provided insight into comprehension. Patterns of errors were examined from the data to gain a more detailed understanding of where participants might go wrong.

Immediate Mental Representation of Probabilities

The data indicated that format influenced accuracy of mental representations; accuracy was lowest for the words format for both younger and older adults. Younger adults were more accurate than older adults across all formats. The age difference was expected; younger adults tend to maintain verbatim or surface representations of text (e.g., Meyer & Pollard, 2006). However, the effect of format on accuracy was not expected. It was hypothesized that older adults would do as well if not better with the words format as verbal knowledge is positively associated with age (e.g., Park et al., 2002); the data indicated otherwise. The data suggested that frequency and percent were more often correctly mentally represented than words by both younger and older adults. These data are consistent with the extant literature (e.g., Cuite, Weinstien, Emmons, & Colditz, 2008).

Both younger and older adults used binary terms such as "might" or "may" when representing probabilities expressed as words. This is inconsistent with the literature that has found that people overestimate verbal expressions of risk (e.g., Knapp, Gardner, Carrigan, Raynor & Woolf, 2009). Instead, the data suggested that younger and older adults used a binary response (e.g., a problem may or may not happen). They were acknowledging that the problem could occur but without any indication of likelihood. This pattern suggests that these younger adults were less exact in representing verbal phrases of probabilities. Older adults had more instances of incorrect mental representations of verbal phrases than younger adults, which is consistent with literature. The high-level binary expressions of risk could be considered gist extraction of the information as would be predicted by the aging literature (e.g., Johnson, 2003; Meyer & Pollard, 2006) and by fuzzy trace theory (Reyna, 2008).

Schapira, Nattinger, and McHorney (2001) found that several participants endorsed that frequencies were easy to understand. The current results indicated otherwise with respect to the incorrect representations provided by six older adults who said, "17 **to** 20" instead of "17 **in** 20." Although the words "to" and "in" are similar, they are not synonymous. Using "to" instead of "in" changes the meaning of the phrase: "17 to 20" people indicates a certain number of people will experience a problem, whereas "17 in 20" people indicates a high likelihood of experiencing a problem. This level of detail regarding such a misunderstanding is an important finding.

Consistent with the recommendation to avoid decimals when communicating risk probabilities (Burkell, 2004; Lipkus, 2007), incorrect representations were given most often by older adults for the 0.01% probability. Six of the seven older adults who incorrectly represented this probability did not include a decimal, making their representation of 0.01% off by a magnitude of 100.

Comprehension of Probabilities

Accuracy on comprehension questions corroborated the immediate mental representation findings: Format influenced accuracy. Performance was worst when probabilities were expressed as words for both younger and older adults. Future research could investigate if training participants to consider verbal expressions of probability in a more quantitative way might support more accurate representations and comprehension. Another avenue of research would be to investigate if explicitly asking participants to elaborate what such a verbal probability expression might mean to encourage deeper processing might support more accurate representations and comprehension (e.g., Natter & Berry, 2005).

Younger adults were also more accurate than older adults; this was not expected as the design was intended to minimize age-related differences in memory as the passages were available for reference while answering questions. This approach provided environmental support for participants (Morrow & Rogers, 2008). Additionally, the reading comprehension literature suggests that age-related differences are minimized when the participants can self-pace (Johnson, 2003) as they could in this task. Older adults also are more likely to have greater experience than younger adults in the health domain, potentially giving them an advantage over younger adults for this task. Moreover, the passages were written at an 8^{th} grade reading level. On the other hand, the task did impose working memory demands and required inferencing, two abilities that show age-related declines (e.g., Park et al., 2002).

Relating Immediate Mental Representation and Comprehension of Probabilities

The correlation between accuracy of immediate mental representation of probabilities and accuracy on comprehension questions was strongly positive for the older adults across all formats, but only for the younger adults in the words format. This difference in relationships between younger and older adults for the frequency and percent formats was likely due to the low number of younger adults (n = 1 and 2, respectively) who were incorrect in their immediate mental representations.

These data indicate the importance of the immediate representation of probabilities and can provide a cue to healthcare providers when performing the teach-back protocol with patients. If a participant does not accurately represent a probability, there is a strong indication that that participant will not achieve high accuracy on a comprehension test. These data are consistent with the extant literature that has identified a significant correlation between teach-back and comprehension (e.g., Fink et al., 2010).

However, Fink and colleagues conducted a between-participants experiment in which half the participants performed the teach-back and the other half did not. The current results are a unique contribution to the teach-back literature as this study drilled down further into the nuances of the teach-back. As all participants performed the teach-back, the data provide a more detailed picture of the errors that participants can make and how such errors relate to comprehension performance. Future studies can investigate various remediation approaches during the teach-back process such that when a patient has an incorrect mental representation, the healthcare provider can focus on rectifying the mistake to improve comprehension.

Delayed Mental Representation of Probabilities

Consistent with fuzzy trace theory (e.g., Reyna, 2008; Reyna, Nelson, Han, & Dieckmann, 2009), the majority of participants did correctly recall the ordinal ranking of problems (i.e., mild > moderate > severe problems) across the three passages. The percent format best supported gist comprehension. Participants were able to retain the gist of the probability information and recognize the correct ordinal rankings.

Unlike the data pattern for immediate mental representations, format did not influence accuracy of delayed mental representations. Younger adults were high in accuracy across all formats; older adults were much worse. Age-related differences were not expected because of the multiple presentations of equivalent probabilities. Participants read three passages each with three expressions of probability in different but equivalent formats, therefore it was expected that participants' recall would be accurate. That is, if participants understood the probabilities, each probability would be reinforced with each presentation despite the format change. Poor cued recall performance could suggest that participants were unable to translate between equivalent formats; they did not understand the probabilities.

These results have provided insight into how people mentally represent and understand an important component of risk, namely, probability. Additionally, the influence of format on mental representation and comprehension of probability was examined as well as the role of age. The patterns of errors participants made have provided insight into how to more effectively communicate probability information for younger and older adults. The results indicate that frequency or percent formats should be used to communicate health risk probabilities to younger and older adults.

CHAPTER 6: PHASE 2 OVERVIEW

Research Questions

The aim of the second phase was to understand how accurately and precisely younger and older adults represented probabilities using a magnitude comparison paradigm. Probabilities were presented as frequencies, percents, and words; age was used as a grouping variable. Participants' accuracy was measured as was response time for accurate trials.

Method Overview

A magnitude comparison paradigm was used in which participants were presented with two probabilities (in the same format) on the screen simultaneously. Participants identified which of the two probabilities was greater. Pairs of probability expressions were classified as near, mid, and far determined by the difference between the probability values of 0.10, 0.25, and 0.40, respectively.

Hypotheses

No differences in accuracy were expected for this task by format or age group. It has been shown that younger and older adults are highly accurate for percent, frequencies, and whole number comparison tasks (e.g., Peters, Slovic, Vastfjall, & Mertz, 2008). Although this is the first study that has used verbal phrases in such a task, accuracy differences were not expected because the verbal phrases were commonly used (e.g., "certain").

It was expected that format would influence performance with respect to response time. Specifically, participants would have fastest response times for percent formats;

frequencies and words would be significantly slower. Comparing percents is essentially a whole number comparison task in which people do well (e.g., Peters et al., 2008). Comparing frequencies with differing denominators requires at least one transformation to complete the task. Comparing verbal phrases of probabilities requires reading longer phrases than the other two formats.

The distance effect (or the precision of one's mental number line) was hypothesized to differ by format, such that the percent format would show the smallest distance effect. It was expected that older adults would be approximately 1.5 times slower overall (e.g., Verhaeghen & Salthouse, 1997), but that their distance effect would be proportionally similar to younger adults across the formats.

CHAPTER 7: PHASE 2 METHOD

The second phase of the study was a magnitude comparison task that investigated the accuracy and precision of mental representations of probabilities as a function of format (percentage, frequency, or words) and age group. Participants determined which of two probability expressions was greater. Accuracy and response times for correct trials were assessed (e.g., Bonato, Fabbri, Umilta, & Zorzi, 2007; Moyer & Landauer, 1967; Peters, Slovic, Vastfjall, & Mertz, 2008).

Method

Participants

Participants were the same as Phase 1. All participants completed Phase 1 and Phase 2 on the same day except for one older adult who returned to the laboratory on another day to finish Phase 2. This participant had a sore leg from a recent fall; sitting was becoming painful by the end of Phase 1. One older adult chose not to participate in Phase 2. The following data are from 36 younger adults and 35 older adults.

Materials

Each participant performed the experimental task on a Dell Dimension 2350 computer with a 17-inch monitor and a standard keyboard. Participants were seated approximately 20 - 25 inches from the monitor. Pink noise of approximately 55 decibels was used to reduce noise distractions during the experimental session; the sound machine was located outside of the experimental room door. Participants were tested individually. E-Prime 2 was used to present the stimuli to the participants (Psychological Software Tools, 2003).

Stimuli

Participants were presented with two probabilities to compare. The stimuli represented values from 0 (impossible) to 1 (certain) in the following formats: frequency, percent, and words. Each probability was presented in white on a black background. The stimuli were presented in sans serif font and were 1 cm in height. One probability expression was presented on the left side of the screen; one was presented on the right side. Width of the stimuli varied as a function of format. Word probability expressions were widest; percent probability expressions were narrowest.

Pairs of probability expressions were determined based on the difference between the pairs: near, mid, and far with differences of 0.10, 0.25, and 0.40, respectively. The verbal phrases for probability were taken from Windschitl and Wells (1996). See Appendix D for their verbal probabilities and numerical equivalents. Table 7.1 shows the stimuli, and Figure 7.1 panels a, b, and c provide examples of the magnitude comparison task for frequency, percent, and words, respectively.

Table 7.1

Stimuli for Magnitude Comparison Task

Frequency	Percent	Words
0 in 1	0%	Impossible
1 in 10	10%	Extremely unlikely
1 in 4	25%	Unlikely
2 in 5	40%	Somewhat unlikely
1 in 2	50%	As likely as is unlikely
3 in 5	60%	Somewhat likely
3 in 4	75%	Likely
9 in 10	90%	Extremely likely
1 in 1	100%	Certain

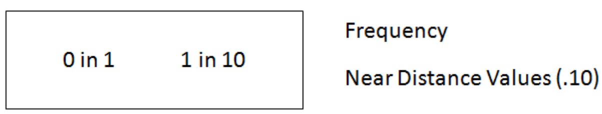

Figure 7.1.a. Example near distance comparison trial for frequency format.

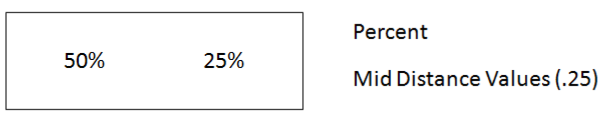

Figure 7.1b. Example mid distance comparison trial for percent format.

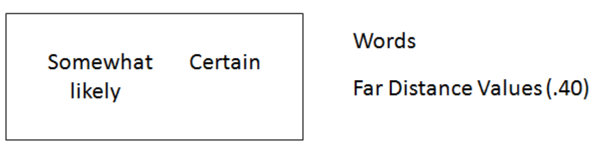

Figure 7.1c. Example far distance comparison trial for words format.

Procedure

A minimum of a five-minute break was given to the participants before Phase 2 started. Figure 7.2 provides an overview of the procedure for Phase 2.

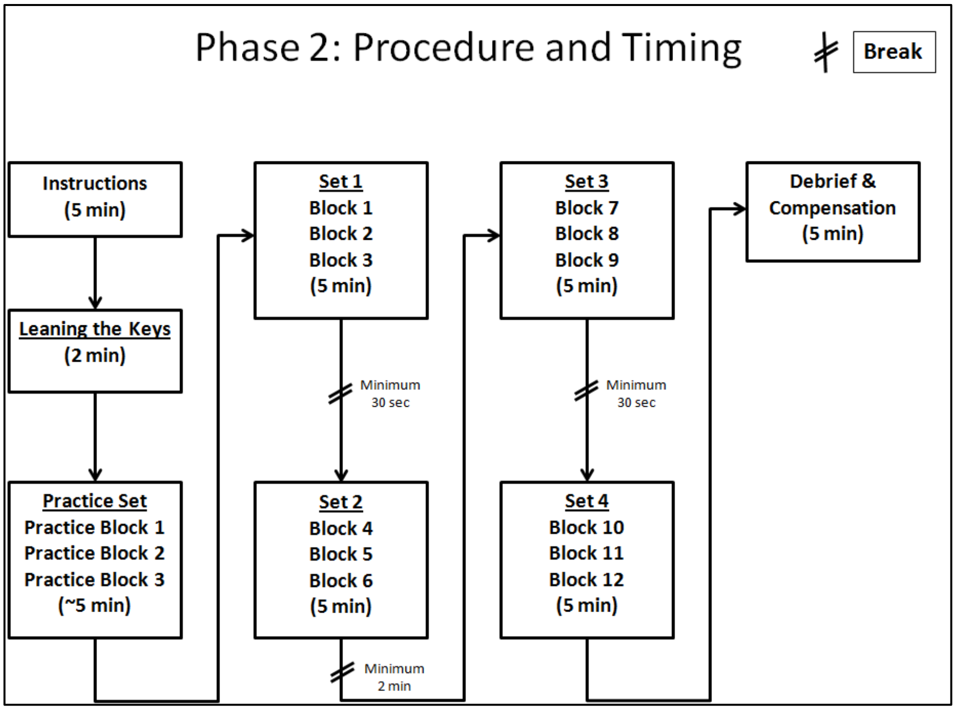

Figure 7.2. Procedure and timing details for Phase 2.

Learning the Keys

Participants completed 24 trials of a stimulus-key mapping task to learn the keys, labeled "Left" and "Right." A white square appeared randomly on either the left or the right side of the screen; the square appeared 12 times on the left and 12 on the right. The participant pressed the key that matched the location of the square.

Practice

Next participants completed practice trials of the experimental task, in which they indicated which of two probabilities presented on the screen was greater. They were instructed to work as quickly and as accurately as they could. There were six trials per practice block divided by format (frequency, percent, and words). The probabilities used for the practice trials were not used for the experimental trials. Each practice block had two trials of near, mid, and far comparisons. Half of the correct answers were "left" and half were "right." The comparisons were randomly presented. Participants received feedback about the accuracy and response time after each trial.

<u>Probability Comparison Task</u>

Participants were asked to identify which of two probability expressions presented on the screen was greater as quickly and as accurately as they could. If the greater probability was on the right side of the screen, the participant used his or her right finger to press the "RIGHT" key (i.e., the "P" key labeled "RIGHT"). If the greater probability was on the left side, the participant used his or her left finger to press the "LEFT" key (i.e., the "W" key labeled "LEFT").

A fixation cross was presented in the middle of the screen for 600 ms, followed by a blank screen for 1000 ms. The probability expressions were presented for 10,000 ms or until the participant responded. This timing was based on pilot testing with younger and older adults; 10 seconds provided ample time to respond but keeps the experiment moving along within a reasonable time frame if there is no response.

There were four sets of three blocks for a total of 288 trials. One block comprised 24 trials; one set comprised three blocks—one for each format of probability expressions (frequency, percent, and words). For example, if a participant was randomly assigned to the words, frequency, percent counterbalance order, that participant would complete one block of 24 trials with words followed by one block of 24 trials with frequency followed by one block of 24 trials with percent. That cycle would repeat three more times for a total of 288 trials.

Format was counterbalanced to control for order effects. See Appendix T for the full counterbalance scheme. The presentation of the stimuli was randomized for each block. Participants could take a break between each block within a set. However, once a block was started, it would not stop until all 24 trials had been completed. A mandatory 30 second break occurred after the first and third sets; a two minute break occurred after the second set.

Before the beginning of each block in sets 2, 3, and 4, participants received feedback based on their performance on the previous block. If a participant missed more than three trials in a block (i.e., accuracy was less than 87.5%), the feedback requested that the participant try to be more accurate. The experimenter encouraged the participant to take more time and try to be more accurate. If a participant did not miss any trials in a block (i.e., accuracy is 100%), the feedback requested that the participant try to be faster. The experimenter encouraged the participant to push himself/herself and try to work faster. If a participant missed one to three trials in a block, the feedback stated, "Great Job! Keep up the good work." The experimenter re-iterated the feedback. Upon completion of the magnitude comparison task, participants were debriefed and compensated for their time.

Design

The experiment was a 3 (Format: percentage, frequency, words) x 3 (Distance: near, mid, far) x 2 (Age: younger adults, older adults) quasi-experimental split plot design. The format and distance variables were manipulated within-subjects; age served as a grouping variable. The dependent variables were accuracy and response time for accurate trials.

CHAPTER 8: PHASE 2 RESULTS

Analysis Overview

Format (frequency, percent, words) and distance (near, mid, far) were manipulated within participant and age served as a grouping variable (younger, older). Alpha was set at .05. Only data from blocks 3 and 4 were included in the following analyses to reflect stable performance. Outlier analyses were conducted using the box plot function in SPSS: Participants with extreme values more than three times the interquartile range were excluded from the analyses.

Accuracy of Probability Comparisons

Participants' accuracy goal was approximately 90%. Younger adults were successful across all conditions, but older adults were only successful for the percent condition. See Figure 8.1. One younger adult was identified as an outlier and excluded from the analysis.

An age x format x distance ANOVA yielded significant main effects for format ($F(2, 67) = 44.37$, $p < .001$, $\eta_p^2 = 0.57$), distance ($F(2, 67) = 24.27$, $p < .001$, $\eta_p^2 = 0.42$), and age ($F(1, 68) = 24.95$, $p < .001$, $\eta_p^2 = 0.27$). The following significant interactions were identified: format x age ($F(2, 67) = 23.45$, $p < .001$, $\eta_p^2 = 0.41$), format x distance ($F(4, 65) = 12.09$, $p < .001$, $\eta_p^2 = 0.43$), and format x distance x age ($F(4, 65) = 2.74$, $p = .036$, $\eta_p^2 = 0.14$).

Pairwise comparisons for the younger adults revealed that percent was most accurate followed by frequency and then words (p's < .05). For the older adults, however, the frequency and words conditions did not differ but both were less accurate than percent (p's $< .05$). See Figure 8.1.

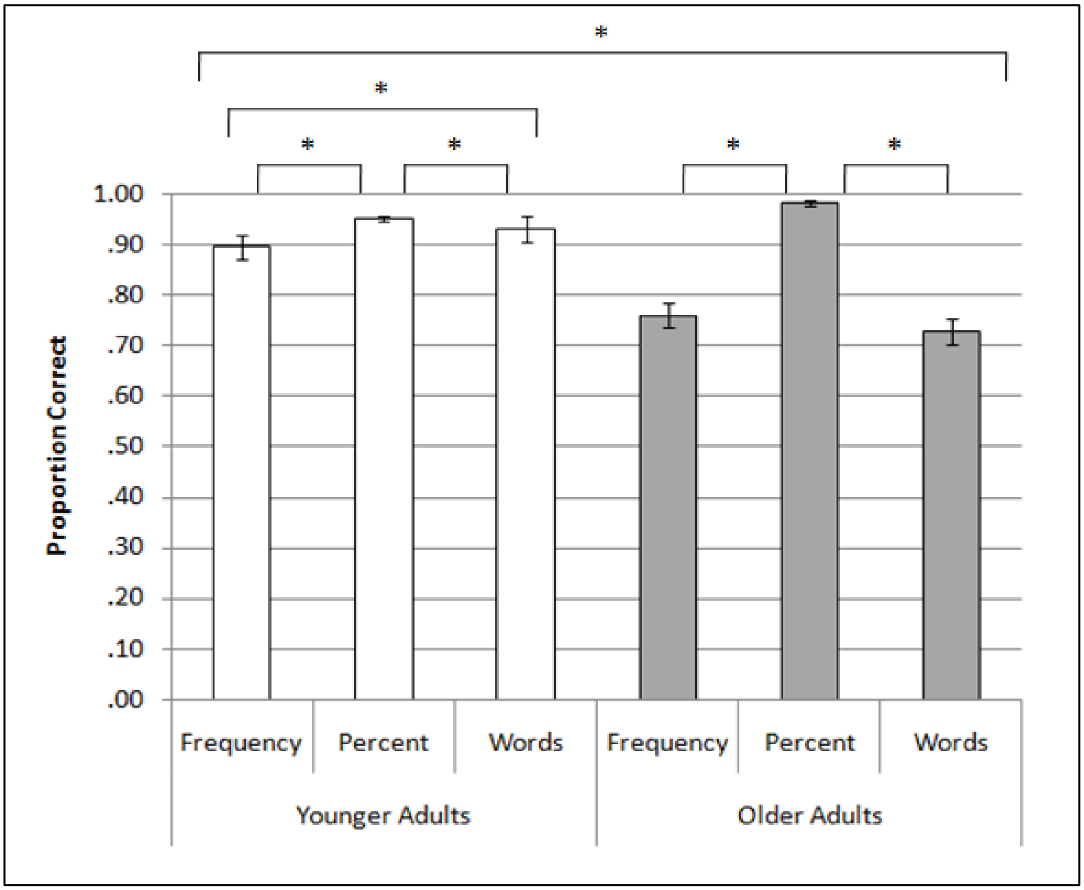

Figure 8.1. Mean accuracy by format and age group. Error bars represent standard error of the mean. *$p < .05$.

Pairwise comparisons for the frequency condition by age group revealed that near distance accuracy was lower than mid and far (p's $< .01$) for both age groups. For the percent condition, no differences in accuracy were identified by distance for younger or older adults. For the words condition, no significant differences were identified for younger adults, but near distance accuracy was lower than mid and far (p's $< .01$) for older adults. See Figure 8.2.

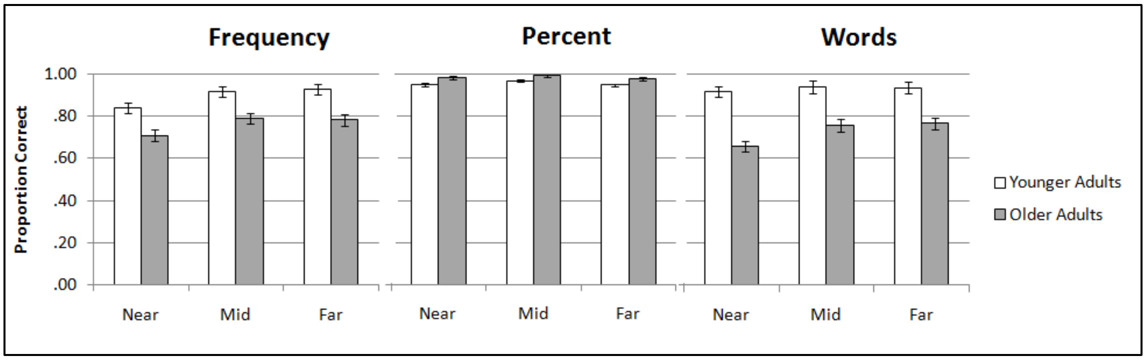

Figure 8.2. Mean accuracy by distance, format, and age. Error bars represent standard error of the mean.

Summary of Accuracy Results for Comparison of Probabilities

The data indicate that format influenced accuracy. Participants had highest accuracy with percent comparison trials. An effect of format on accuracy was not expected. A main effect of distance was identified; participants were least accurate for near distance comparisons.

The data show that younger adults were more accurate than older adults. A difference in accuracy between the age groups was not expected. Participants had ample time (up to 10 seconds) to make comparisons; they were told to work as quickly and as accurately as they could. In a similar task with older and younger adults, Peters and colleagues did not find an age difference in accuracy (Peters, Slovic, Vastfjall, & Mertz, 2008).

The significant format by age interaction indicated that younger and older adults were most accurate with percent. Younger adults were less accurate with words and frequency, as were older adults. However, older adults showed a 20% drop in accuracy from percent to frequency and words trials.

Precision of Mental Representations of Probabilities

The goal of this analysis was to investigate the precision of mental representations of probabilities using the distance effect; the smaller the difference in response times between near and far distance trials, the more precise mental representation. Only response times for correct trials were used, and only data from blocks 3 and 4 were included in the analyses to reflect stable performance.

Each format was analyzed separately to maximize the number of participants per analysis and thus the power of the analyses. Outlier analyses identified three younger adults and five older adults that were excluded from the frequency and from the words formats. No outliers were identified for the percent format. It was also assumed that response times would differ by format as the amount of information required to process for each format differed such that percent would be fastest and words would be slowest. Only near and far trials were used to create a linear slope for the distance effect.

Distance x age ANOVAs across all formats revealed a main effect of distance such that near trial response times were always greater than far trial response times: The distance effect was identified for all formats and ages. A main effect of age was also identified across all formats such that younger adults were always faster than older adults. A significant distance x age interaction was identified for the frequency format only ($F(1, 61) = 8.55$, $p = .005$, $\eta_p^2 = 0.12$); there was a greater difference between RTs for near and far trials for older adults than younger adults. See Table 8.1 for the statistics and Figure 8.3 for an illustration of mean performance by format, distance, and age.

Table 8.1

F-Statistics by Format for Distance and Age

Frequency	
Distance	$F(1, 61) = 45.65$, $p < .001$, $\eta_p^2 = 0.43$
Age	$F(1, 61) = 37.87$, $p < .001$, $\eta_p^2 = 0.38$
Distance x Age	$F(1, 61) = 8.55$, $p = .005$, $\eta_p^2 = 0.12$
Percent	
Distance	$F(1, 69) = 5.09$, $p = .027$, $\eta_p^2 = 0.07$
Age	$F(1, 69) = 120.96$, $p < .001$, $\eta_p^2 = 0.64$
Distance x Age	$F(1, 69) = 0.47$, $p = .50$, $\eta_p^2 = 0.01$
Words	
Distance	$F(1, 61) = 7.41$, $p = .008$, $\eta_p^2 = 0.11$
Age	$F(1, 61) = 29.67$, $p < .001$, $\eta_p^2 = 0.33$
Distance x Age	$F(1, 61) = 0.00$, $p = .99$, $\eta_p^2 = 0.00$

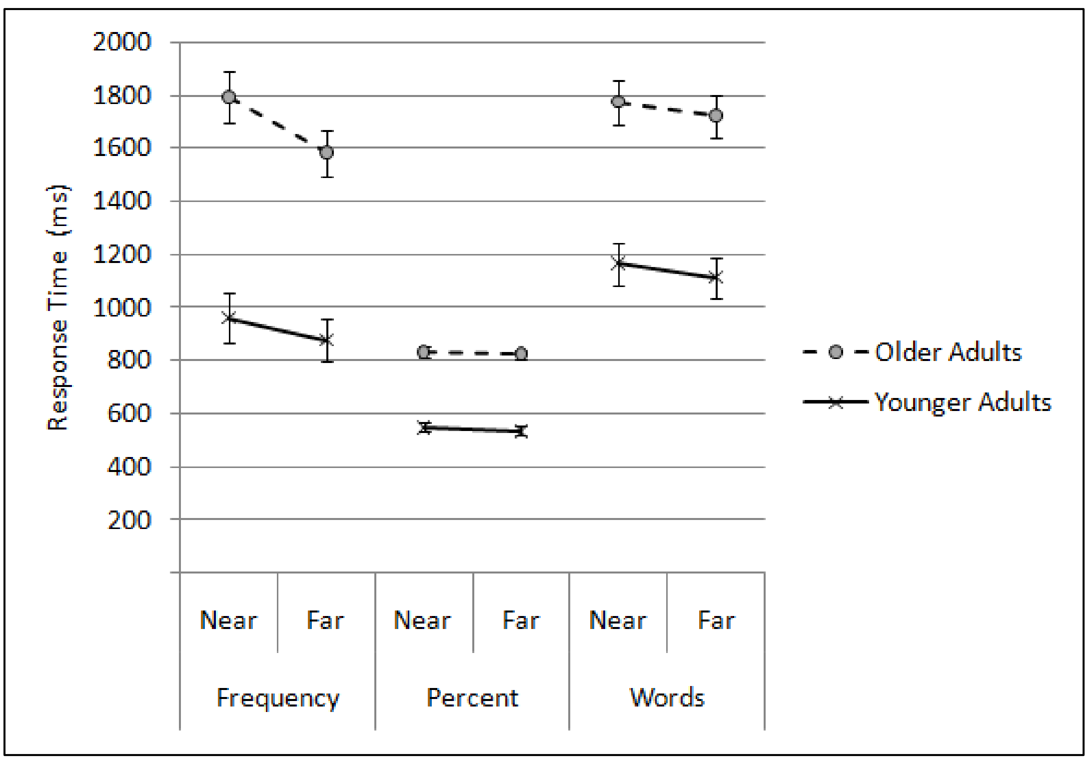

Figure 8.3. Mean response times by format, distance, and age for correct trials only. Error bars represent standard error of the mean.

As the presence of the distance effect for all formats was confirmed, a format x age ANOVA was conducted using the slope (or difference between near and far trial RTs) as the primary dependent variable to investigate precision of mental representation by format and age. A total of 30 younger adults and 26 older adults had complete data for all three formats after outliers were excluded. A significant main effect of format ($F(2, 53) = 19.35$, $p < .001$, $\eta_p^2 = 0.42$) and interaction of format x age ($F(2, 53) = 3.65$, $p = .033$, $\eta_p^2 = 0.12$) were identified.

Pairwise comparisons by format revealed that the percent slope was smallest (most precise) and the frequency slope was largest (least precise; all *p*'s < .05). The interaction of format x age was driven by the older adults' much larger slope for the frequency format than younger adults' slope. See Figure 8.4 for an illustration of mean slopes by format and age.

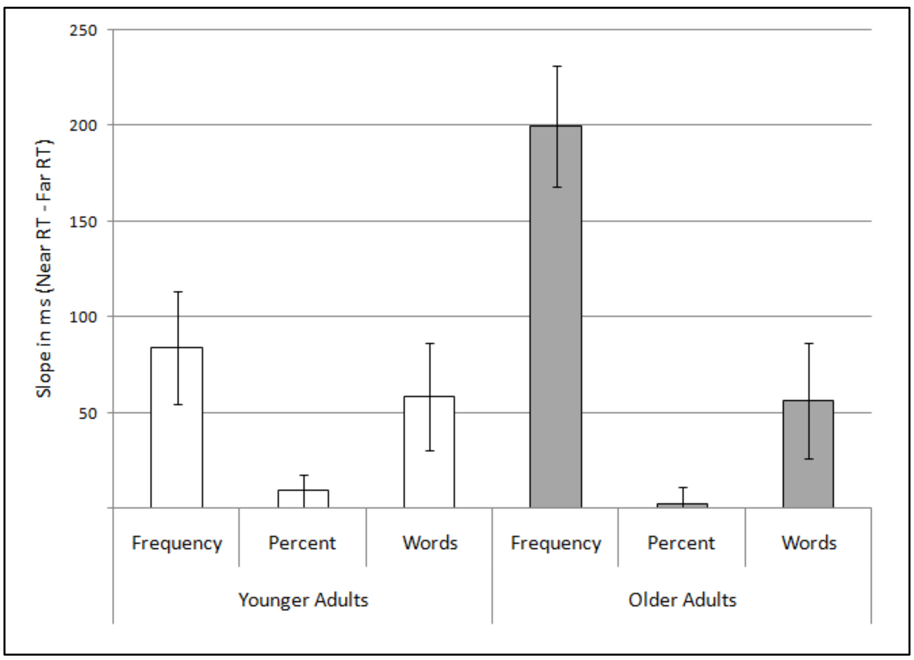

Figure 8.4. Mean differences in response times (slopes) by format, distance, and age for correct trials only. Error bars represent standard error of the mean.

Summary of Precision of Mental Representations of Probabilities

Across all formats, a similar pattern emerged in which significant effects of distance and age were identified. As expected, near distance comparisons took longer than far distance comparisons, consistent with the distance-effect (e.g., Dehaene, Bossini, & Giraux, 1993). As expected, older adults were slower than younger adults across all formats. Age-related differences in response times for the percent and words comparisons were consistent with the literature: Older adults were 1.5 times slower than younger adults (e.g., Verhaeghen & Salthouse, 1997). Older adults were 1.8 times slower in the frequency condition, which was more than expected.

Analysis of the slopes revealed that precision of mental representation of probabilities is influenced by format. Percent format was mentally represented most precisely by all participants, whereas frequency format was represented least precisely.

CHAPTER 9: PHASE 2 DISCUSSION

Using a magnitude comparison task provided an alternative approach to investigating participants' mental representations of probabilities expressed as frequencies, percents, and words. This is the first study that has used verbal expressions of probability in a magnitude comparison task. One goal was to investigate how accurately verbal expressions were compared and if qualitative expressions of numerical information were mentally represented in a similar spatial format akin to the mental number line. Results demonstrated that the percent format led to highest accuracy and had the most precise mental representation for younger and older adults. Word comparisons did show a distance effect. Older adults' accuracy was significantly lower than younger adults for the frequency and words conditions.

Accuracy of Mental Representations for Probability Comparisons

Older adults' low accuracy for frequency comparisons was consistent with the results of a fraction comparison study in which university and junior college students participated (Schneider & Siegler, 2010). Fraction and frequency comparisons are very similar in that they both represent a portion of the whole. The mean error rate was 6% for the university students and 30% for the junior college participants; the response times were much higher for the junior college students. However, the data patterns were the same between the groups such that a distance effect was observed. The older adult data for the frequency format were consistent with the junior college participants. In fact, the mean error rate for the older adults was slightly less at 25%.

A potential explanation for the age difference in accuracy for the frequency trials might be extrapolated from DeWolf and Vosniadou (2011), who also investigated the mental representations of fractions. They classified fraction comparison pairs as "consistent" and "inconsistent." Consistent fraction pairs were those with larger numbers representing the larger fraction value (e.g., 7/8 compared to 1/2). Inconsistent fraction pairs were those with the smaller number values for the numerator and the denominator representing the larger fraction values (e.g., 2/3 compared to 3/8). Participants were much more accurate with consistent comparisons. They proposed that participants had to inhibit their whole number bias when comparing fractions that were inconsistent. Research has suggested that older adults have difficulty inhibiting processes (e.g., Hasher, Stoltzfus, Zacks, & Rypma, 1991). Thus, having to inhibit a propensity for a whole number bias might have put older adults at a disadvantage for the frequency comparison task resulting in lower accuracy than younger adults.

The stimuli used in this study were balanced across consistent and inconsistent trials; future research should include older adults in a study that focuses on the effect of consistent and inconsistent fraction comparisons. Additionally, fuzzy trace theory predicts gist extraction of probabilities such that only numerator information extracted, resulting in denominator neglect (e.g., Reyna et al., 2009) potentially explaining the older adults' low accuracy.

Older adults did not experience an advantage when working with verbal expressions of probabilities; in fact, older adults' mean error rate was 28% for the words comparison task. Comparing verbal phrases might have posed difficulties for older adults because the instructions stated to identify the greater probability: When participants were comparing verbal expressions representing less than 50% they essentially had to process a double-negative. For example, comparing "unlikely" to "impossible" might have been difficult for participants to recognize that "unlikely" represents a greater probability than "impossible."

These results suggest that percent (i.e., whole numbers) should be used as such values are accessed accurately and quickly. This is inconsistent with the recommendations of Gigerenzer and colleagues (e.g., Gigerenzer et al., 2008; Gigerenzer & Edwards, 2003). However, the primary focus of Gigerenzer's arguments for use of natural frequencies is for Bayesian tasks; not for probability comparisons. Thus, the results of this study have defined a boundary condition for the natural frequency proposal: Frequencies are not appropriate for all tasks, specifically probability comparisons.

Precision of Mental Representations of Probabilities

The distance effect slopes for the percent and word formats were consistent with past findings from Geary and Lin (1998) and Peters, Slovic, Vastfjall, and Mertz (2008) such that there were no differences in slopes by age. However, the current study identified a format by age interaction such that older adults had a much steeper slope for the frequency format than did younger adults. Although Peters and colleagues included percent and frequency formats, their frequency stimuli always used a denominator of "100." This essentially created a whole number comparison task of the numerators, whereas in the current study, frequency comparisons required accessing the real value of the fraction. Older adults' distance effect slope suggests that they were differentially impacted by the frequency format; as discussed earlier, it could be that they struggled inhibiting their whole number bias when comparing fractions that were inconsistent.

The data pattern for words indicated a distance effect: Both younger and older adults were faster on average to compare distant trials than near trials. This might indicate that verbal expressions are represented spatially on a mental number line. More research is needed to focus on the qualitative probability expressions and understand how such information is mentally represented.

Interestingly, Schneider and Siegler (2010) noted that neuroimaging studies have identified the intraparietal sulcus as active during spatial and numerical tasks (e.g., Tudusciuc & Nieder, 2007). As spatial abilities are susceptible to age-related declines (e.g., Park et al., 2002), a future avenue for research should investigate the extent to which spatial ability can predict the precision of mental representations using the distance effect for various formats of number and probability comparison tasks.

CHAPTER 10: PHASE 3 OVERVIEW

Research Question

The goal of the third phase of the study was to examine the relationships between format, age, numeracy, comprehension, and mental representations. The third phase combined the results of the first two phases to provide an overall picture of how probabilities were mentally represented and comprehended as a function of format, age, and numeracy.

Correlations and hierarchical multiple regression analyses were performed to examine the relationships between various participant characteristics and multiple dependent variables including: comprehension question accuracy, probability comparison accuracy, and precision of mental representations of probabilities (i.e., distance effect for the probability comparison task). Numeracy was included as a continuous variable; age was included as a categorical variable.

Hypotheses

It was expected that numeracy would be positively related to accuracy on comprehension tests of the Phase 1 study irrespective of format. Additionally, it was expected that numeracy would predict precision of mental representations. The fact that performance can be influenced by probability format suggests that mental representations of probability differ in precision, especially for lower numerate people (Peters et al., 2006). It was expected that the relationship between accuracy on comprehension questions (Phase 1) and accuracy on probability comparison tasks (Phase 2) would be highly correlated.

CHAPTER 11: PHASE 3 METHOD

Method

The third phase of the study examined the relationships between numeracy and Phase 1 and 2 performance. The same participants were in each study making this analysis possible, except for one older adult who chose not to participate in Phase 2. A hierarchical multiple regression analysis was conducted to examine the proportion of the variance accounted for by numeracy to predict comprehension.

CHAPTER 12: PHASE 3 RESULTS

Ability Tests

Several ability tests were administered during Phase 1 to assess a range of cognitive abilities; both general and health-specific tests were given (refer to Table 3.2 for details and references). Table 12.1 shows the mean performance for each ability test by age group; a significant difference between age groups is indicated by an asterisk.

Table 12.1

Ability Test Data for Participants

	Younger Adults (*n*=36)		Older Adults (*n*=36)		
Ability Test	*M*	*SD*	*M*	*SD*	t-value
Reading Comprehension	9.8	0.6	8.1	2.0	4.9*
Reading Rate	300.3	83.6	274.2	81.6	1.3
S-TOFHLA	35.5	0.7	34.3	3.0	2.4*
TOFHLA-Numeracy	15.6	1.0	14.9	1.8	1.9
REALM-SF	6.9	0.3	6.9	0.4	0.00
Numeracy	11.2	1.0	7.3	3.3	6.6*
Vocabulary	31.9	3.0	33.5	5.4	-1.5
Reverse Digit Span	9.1	2.2	7.4	2.8	2.9*
Subtraction & Multiplication	23.3	9.4	23.1	9.2	0.02
Digit-Symbol Substitution	72.8	10.4	52.8	11.1	7.9*
Math Anxiety	50.5	16.8	57.4	23.5	-1.4

*$p < .05$.

Older adults' numeracy test performance was slightly lower compared to older adults' performance in previous studies (Donelle, Hoffman-Goetz, & Arocha, 2007; Galesic, Garcia-Retamero, & Gigerenzer, 2009; Peters et al., 2008), whereas younger adults' numeracy accuracy was higher compared to previous studies (Galesic et al; Peters et al.). Age-related differences in vocabulary, memory span, and speed of processing are often observed (e.g. Rogers, Hertzog, & Fisk, 2000).

Relating Numeracy to Comprehension and Mental Representation

The literature has indicated that numeracy is related to performance across different number related tasks (e.g., Donelle et al., 2007; Galesic et al., 2009; Peters et al., 2008). The goal of the following analyses was to examine the relationship between numeracy and comprehension accuracy, comparison task accuracy, and precision of mental representations.

Numeracy and Comprehension Questions

Figure 12.1 shows the relationship between number correct on the comprehension questions (30 questions collapsed across format) and numeracy. A strong relationship between these variables was observed for older adults; only a weak correlation existed for younger adults. Numeracy showed a stronger relationship with comprehension question accuracy for older adults ($r(36) = .82$, $p < .001$) than for younger adults ($r(36) = .20$, $p = .25$).

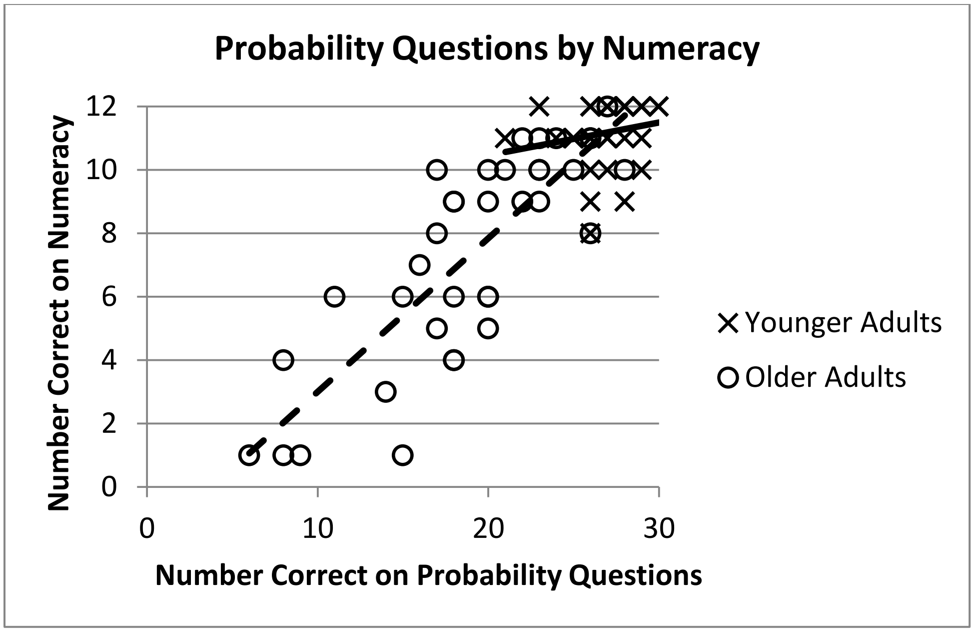

Figure 12.1. The relationship between number correct on the probability comprehension questions and numeracy for younger and older adults.

Numeracy and Comparison Task Accuracy

Table 12.2 provides the correlation coefficients for the relationship between numeracy and comparison task accuracy by format and age. Only data from Blocks 3 and 4 were used to reflect stable performance. A strong relationship between numeracy and comparison task accuracy was observed for older adults with the frequency and words formats. The data suggest no correlation between accuracy on percent comparison trials and numeracy for younger or older adults.

Table 12.2

Correlation between Numeracy and Comparison Task Accuracy by Format and Age

	Frequency	Percent	Words
Younger Adults	$r(35) = .29$	$r(36) = -.04$	$r(36) = .20$
Older Adults	$r(35) = .69$*	$r(35) = .18$	$r(35) = .63$*

*p<.05

Numeracy and Precision of Representations (Distance Effect)

Table 12.3 provides the correlation coefficients for the relationship between numeracy and distance effect by format and age. Only data from Blocks 3 and 4 were used to reflect stable performance. Negligible relationships between numeracy and the distance effect across all formats were observed for all younger and older adults. These results are inconsistent with Peters and colleagues (2008) who found a significant correlation between numeracy and precision of representation.

Table 12.3

Correlation between Numeracy and Distance Effect by Format and Age

	Frequency	Percent	Words
Younger Adults	$r(33) = -.02$	$r(36) = -.15$	$r(33) = .24$
Older Adults	$r(30) = .04$	$r(35) = .02$	$r(30) = .28$

Hierarchical Regression for Accuracy on Comprehension Questions

The goal of the following hierarchical regression was to determine the unique proportion of variance accounted for by numeracy. Based on the literature there are other independent variables that likely predict comprehension. Following approaches used by Morrow et al. (2005), Pak, Czaja, Sharit, Rogers, and Fisk (2008), and Peters et al. (2008), “the logic of the regression was to examine the predictability of chronological age before and after age-related differences in abilities were controlled for. If after controlling for ability differences age is no longer a significant predictor of performance, then the implication is that differences in performance can be explained by differences in those abilities that are known to be age-related” (Pak et al., p. 3048). The goal of this regression was to understand the unique proportion of variance accounted for by numeracy above and beyond that captured by age, general cognitive abilities, and health literacy. Predictor variables were entered with accuracy on comprehension questions collapsed across format as the dependent variable. Table 12.4 shows the hierarchical regression for accuracy on comprehension questions with age entered first, general abilities next, health literacy third, and health numeracy fourth.

When chronological age alone was entered into the regression (Model 1), it accounted for 46% of the age-related variance in comprehension scores. Subsequent regressions controlled for cognitive abilities. Model 2 shows that when controlling for a range of cognitive abilities, the predictability of age was slightly reduced but not eliminated; reading comprehension, vocabulary, and subtraction-multiplication performance accounted for unique variance. In Model 3, I entered health literacy measures; they did not contribute to explain any additional variance. Lastly, in Model 4, I included numeracy. As the addition of numeracy led to a statistically significant increase in R-squared, I have evidence that numeracy predicts comprehension over and above the previously entered independent variables. Age, however, still accounted for a large proportion of the variance.

When the regression was conducted with age entered last to investigate the extent to which age predicts comprehension over and above the previously mentioned variables, there was still a significant change in the model suggesting age-related variance has not been completely accounted for.

Table 12.4

Hierarchical Regression for Accuracy on Comprehension Questions

	Model							
	1		2		3		4	
	B	Change R^2	*B*	Change R^2	*B*	Change R^2	*B*	Change R^2
Age	-.68*	.46*	-.65*	.35*	-.65*	.003	-.37*	.08*
Reading Comprehension			.36*		.39*		.36*	
Reading Rate			.002		.004		.01	
Vocabulary			.27*		..30*		.11	
Reverse Digit Span			.02		.02		-.02	
Digit-Symbol Substitution			-.15		-.16		-.17*	
Subtraction-Multiplication			.22*		.22*		.15*	
Math Anxiety			-.01		-.03		-.03	
S-TOFHLA					-.05		-.07	
REALM-SF					-.04		-.03	
TOFHLA-Numeracy							.11	
Numeracy							.42*	

Note: N = 72.
**p* < .05

These data indicate that age was a significant predictor of performance for accuracy on comprehension questions. After controlling for a range of cognitive variables, reading comprehension, speed of processing (digit-symbol substitution task), math ability (subtraction-multiplication test), and numeracy were significant predictors of comprehension.

CHAPTER 13: PHASE 3 DISCUSSION

These results were consistent with the literature that has identified the role of numeracy as an important factor that is strongly related to comprehension and mental representation of probabilities (e.g., Peters, Slovic, Vastfjall & Mertz, 2008). Numeracy was strongly related to comprehension question accuracy and probability comparison accuracy for older adults. Weaker relationships were observed for the younger adults. The data indicated that younger adults were range restricted on the numeracy test likely weakening the relationship. Future studies should include a more diverse sample of younger adults to further explore the relationships.

Contrary to the results of Peters et al., no relationship was identified in the current study between numeracy and precision of mental representation. They collapsed their data across age and format; they did not include a words condition.

In the hierarchical regression, numeracy was a significant predictor of comprehension on probability questions after controlling for a wide range of cognitive variables. The effect of age was reduced but not eliminated and reading comprehension, speed of processing and math ability contributed unique variance to the model. These were exploratory analyses and the overall approach was not designed to be an individual difference study. Therefore, a limitation of these results is that not all relevant cognitive abilities were assessed nor were multiple measures of each construct measured. Moreover, the current sample's generalizability was limited due to the fact that all younger adults were university students and the older adults were healthy volunteers.

CHAPTER 14: GENERAL DISCUSSION

As healthcare moves toward a shared decision making approach in which patients are more active participants in their own healthcare, it is critical that patients are provided with information they can understand and use when making healthcare decisions. One type of healthcare decision includes comparing probabilities of risks associated with various treatment options. To make informed decisions, probabilities must be correctly understood (e.g., Finucane & Gullion, 2010). However, probabilities are a difficult concept to understand (e.g., Reyna & Brainerd, 2008) and there has been little investigation into how people comprehend health risk probabilities.

Lipkus and Peters (2009) proposed a high-level descriptive theoretical framework of numeracy in health decision-making in which they highlighted comprehension of numerical information as essential to informed decision making (see Figure 14.1). In their framework, Lipkus and Peters identified several factors that influence comprehension of numerical information. The present study focused on the following factors: format of the numeric stimuli (specifically probabilities), mental representation of number magnitude, and numeracy. Additionally, age was included as a grouping variable to investigate how age-related differences in cognitive abilities might influence mental representation and comprehension of health risk probabilities.

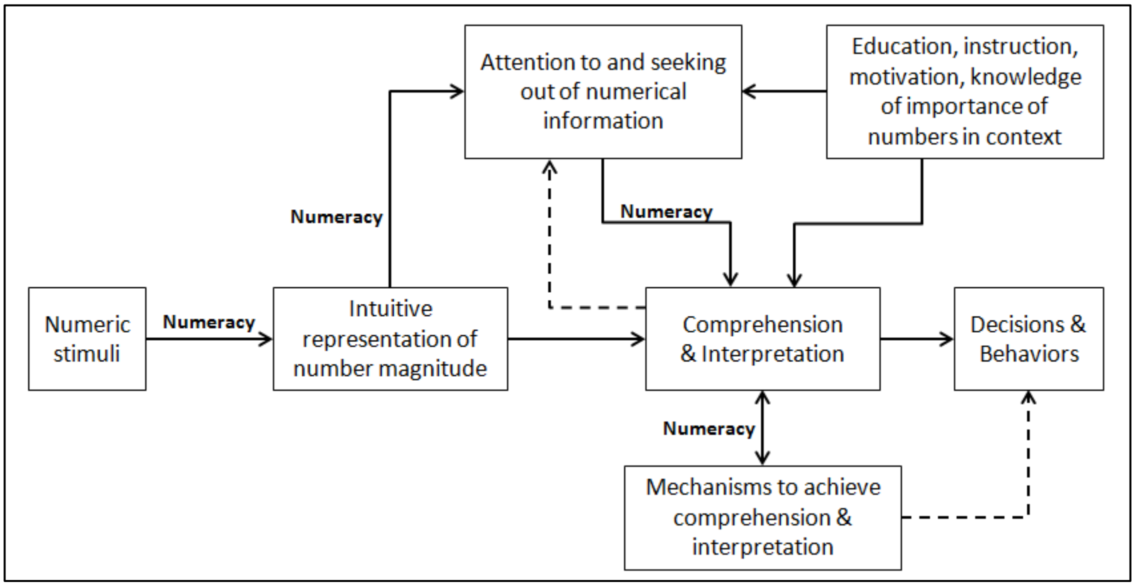

Figure 14.1. Theoretical framework of numeracy in health decision-making (Lipkus & Peters, 2009, p. 1073).

The current research contributed to the numeracy framework in several aspects. Phase 1 provided health risk probabilities to participants embedded within representative health expository texts. Even though participants were not prompted or directed toward the health risk probabilities, nearly all health risk probabilities across all formats were acknowledged in their immediate representations (teach-back responses). This suggests that people do attend to health risk probabilities. Yet, what cannot be discerned from this study is the depth of processing in which each person engaged to understand the probabilities. According to Lipkus and Peters (2009), it would be hypothesized that a person's level of numeracy would influence the depth of processing such that the higher the numeracy, the deeper the processing.

However, it might also be that depth of processing interacts with probability format (Lipkus & Peters). Partial evidence for this comes from participants' teach-back responses in Phase 1 of this study. The younger adults were all high in numeracy and many participants' immediate representations of the verbal expressions of "very likely" were vague, suggesting a cursory or shallow processing of the verbal information similar to the older adults who had a wide range of numeracy.

Additionally, age has been associated with gist rather than verbatim extraction in the reading comprehension literature (for reviews see Johnson, 2003; Meyer & Pollard, 2006). As this task used expository texts within which probabilities were embedded, gist or verbatim mental representations could be inferred from the teach-back responses. Such a difference in information processing might have an impact on the mental representation and comprehension of probabilities. Future studies should investigate and disentangle the factors of processing depth and gist extraction of health risk probabilities by format, age, and numeracy. One avenue of research to explore could be to explicitly manipulate the depth of processing by asking participants to draw pictures or graphs of probabilities or having them elaborate and describe in more detail the meaning of probabilities (Natter & Berry, 2005).

I have revised the numeracy framework to reflect my interpretation of the current study's results; see Figure 14.2. Specifically, "numeric stimuli" is a broad term that encompasses any stimulus that expresses numeric information (e.g., numbers, tables, graphs). The focus of this study was on health risk probabilities presented as frequencies, percents, or words. The inclusion of words, or qualitative expressions of numeric information, is a contribution to the framework as comprehension of verbal expressions of probabilities has received little attention in the literature with respect to mental representation, comprehension, and numeracy.

Contrary to the Lipkus & Peters (2009) proposed framework, a significant relationship between numeracy and precision of mental representations (i.e., distance effect slope) was not identified. The current study did reveal a significant positive correlation between numeracy and immediate mental representation (teach-back response). These data suggest that mental representation should be sub-divided into more specific categories (or boxes) in the framework: precision of mental representation (distance effect slope) and accuracy of mental representation (accuracy of immediate representation). Future research can investigate mental representations of probabilities by employing a task in which participants indicate spatially where they "visualize" various probabilities (across formats) on a mental number line (Nees & Walker, 2011).

Task must also be delineated in the model. The problem space of this study was a health risk probability comparison task. The outcomes and recommendations would likely change depending on the task. For example, the current results suggest that percent format best supports comprehension of health risk probabilities within a comprehension and comparison task space. However, Gigerenzer and colleagues have identified frequencies as best supporting comprehension of Bayesian reasoning problems. Indeed, Lipkus and Peters (2009) acknowledged that the effects of numeracy might vary as a function of task.

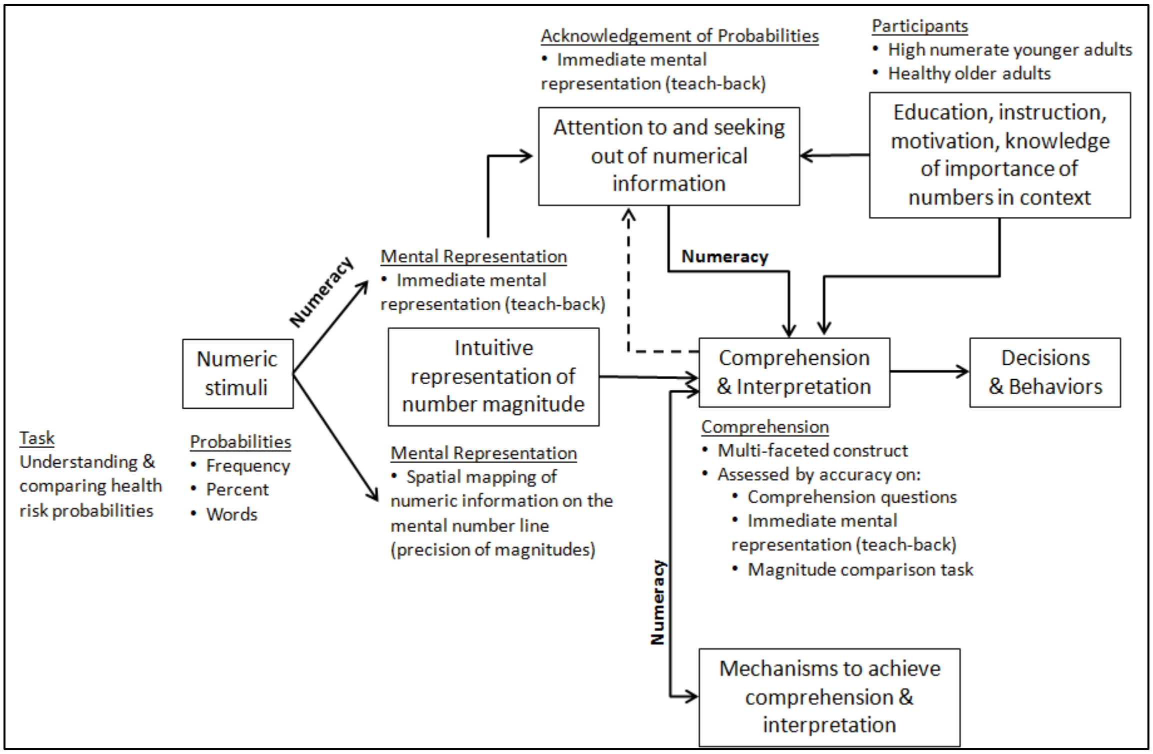

Figure 14.2. Revised theoretical framework of numeracy in health decision-making based on the current findings.

The results from the present study have practical implications for health risk communication. When conveying options for treatments and the associated health risk probabilities to patients (a probability comparison task), percent format will likely best support comprehension of probabilities of 1% or greater for highly numerate younger adults and healthy older adults. More research is needed to investigate optimal format for comprehension when comparing small values describing probabilities of less than 1%.

The detailed analysis of participants' immediate mental representations and the positive relationship with comprehension question accuracy supports the continued efforts to improve patient-provider communication by providing details about specific patient statements that indicate misunderstanding. For example, both younger and older adults used phrases such as, “might” or “could occur” to express verbal probabilities. In practice, healthcare providers can ask patients to elaborate if they use such terms. For the frequency format, some older adults said, “17 to 20,” instead of “17 in 20;” such a subtle change in words significantly changes the expression. Healthcare providers’ time constraints and knowledge about the topic might create a situation of top-down bias influencing what they hear patients say. Training providers about such subtle words changes would likely improve the teach-back process and patient comprehension.

A limitation of the current study is the generalizability of the results to the general population. Specifically, all of the younger adults were students at a university that selects for math and science ability. The restriction of range for numeracy observed in this sample of younger adults would likely not occur if a more diverse sample were included. This is certainly a next step: Including younger adult participants with diverse backgrounds, such as students from junior colleges, liberal arts majors, and those young adults who have a high school degree or less. Schneider and Siegler (2010) identified differences in performance between Carnegie Mellon University students and students from a junior college.

Similarly, research should include a more diverse sample of older adults to increase the generalizability of the results. The older adult participants in this study represented healthy older adults but not older adults who might be underserved, poor, low educated or with worse health. Interestingly, the healthy older adult participants in this study had a wide range of numeracy; it is likely that including more diverse and less healthy older adults that an even wider range of numeracy and performance will be observed. However, the role of experience and familiarity should be accounted for in future studies. Such knowledge might attenuate the role of numeracy in comprehension of health risk probabilities. Researchers should also consider including older patients who are making real-time decisions about their health: How do they comprehend health risk probabilities and how does that correlate or predict decisions made?

Additionally, this study was a cross-sectional extreme age groups design in which cohort effects cannot be ruled out as alternative explanations. As previously mentioned, a cohort effect in arithmetic has been identified (Schaie, 1996) such that comparable results between younger and older adults in precision of mental representation of magnitudes might be due to an educational "buffer" of older adults such that age-related declines make them appear as younger adults. Thus the possibility that there are age-related declines in precision of mental representations cannot be rejected.

A group difference approach was used in Phases 1 and 2 of this study, whereas an individual difference approach was used in Phase 3. The power needed for Phase 3 was likely inadequate for an individual difference study based on the total number of participants (n=72; 36 per age group). However, some interesting relationships still emerged for the older adults with respect to numeracy and comprehension and accuracy on comparison tasks. Future research investigating the role of numeracy as an individual difference variable should include more participants who represent a wide range of education and abilities. Tests of working memory, visuospatial ability, and attention should be included and, each construct should be represented by more than one test.

Although more research is required to understand the factors that influence comprehension of health risk probabilities, this work has contributed to the literature both theoretically and practically. This research has provided empirical evidence that has added more detail to the descriptive framework of numeracy in health decision-making proposed by Lipkus and Peters (2009), and I have proposed a revision of that framework based on the current study's findings (see Figure 14.2). Future paths of research have been identified to further clarify the picture of numeracy, mental representation, format, and age with respect to comprehension of health risk probabilities.

APPENDIX A

EVERYDAY MATH ANXIETY QUESTIONNAIRE

The items in the questionnaire refer to experiences that may cause tension, apprehension, or anxiety. For each item, mark the response that describes <u>how anxious it would make you</u>.

1. Being given a set of multiplication problems to solve on paper.

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

2. Listening to a doctor talk about the chance of a side effect using words (e.g., "It rarely happens.").

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

3. Being given a set of subtraction problems to solve on paper.

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

4. Calculating a 10% discount on an item in a store.

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

5. Knowing how much medicine to take if your doctor tells you to cut back the dose by 1/3.

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

6. Changing a recipe to make half the number of servings.

Not at all	Slightly	Moderately	Very	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

7. Calculating the number of calories eaten for dinner.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

8. Listening to the weather forecaster talk about the chance for rain using frequencies (e.g., "There is a 1 in 5 chance of rain today.").

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

9. Helping a high school student with his or her math homework.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

10. Using exact change to pay for something at the grocery store.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

11. Helping a middle school student with his or her math homework.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

12. Choosing your best option from several different insurance plans.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

13. Balancing your checkbook.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

14. Determining how many doses of cough medicine can be taken in a 24-hour period.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

15. Listening to a doctor talk about the chance of a side effect using frequencies (e.g., “There is a 1 in 100 chance.”).

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

16. Being given a set of addition problems to solve on paper.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

17. Determining if you have enough money for a bus fare.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

18. Reading a cash register receipt after you buy something.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

19. Helping an elementary student with his or her math homework.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

20. Listening to a pharmacist tell you how much medicine to take.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

21. Determining the amount of detergent to use for a half load of laundry.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

22. Listening to the weather forecaster talk about the chance for rain using words (e.g., “There is a small chance of rain today.”).

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

23. Making a weekly grocery budget.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

24. Knowing how much medicine to take if your doctor tells you to cut back your dose by 33%.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

25. Being given a set of division problems to solve on paper.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

26. Determining how many calories are in a food item by looking at the nutrition label.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

27. Changing a recipe to double the number of servings.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

28. Finding the best deal on a product by comparing discount prices.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

29. Listening to a doctor talk about the chance of a side effect using percentages (e.g., “There is a 1% chance.”).

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

30. Determining a food serving size for half the calories per serving.

Not at all	Slightly	Moderately	Very	Extremely
☐$_1$	☐$_2$	☐$_3$	☐$_4$	☐$_5$

31. Choosing a low calorie food by reading the nutrition label.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

32. Making a monthly budget.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

33. Listening to the weather forecaster talk about the chance for rain using percentages (e.g., "There is a 20% chance of rain today.").

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

34. Calculating how much money is needed for a taxi when given the cost per mile.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

35. Calculating a 15% tip.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

36. Using numerical information.

Not at all	Slightly	Moderately	Very	Extremely
☐1	☐2	☐3	☐4	☐5

APPENDIX B

DEVELOPMENT OF HEALTH PASSAGES

Probability Format: PERCENT

Word Count							Flesch-Kincaid Grade Level					
Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Range**	**Difference (Max-Min)**	Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Difference (Max-Min)**
1	25	27	26	26.0	25-27	2	1	10.0	9.7	9.5	9.7	0.5
2	28	25	28	27.0	25-28	3	2	12.0	11.1	12.0	11.7	0.9
3	30	29	33	30.7	29-33	4	3	5.2	5.6	5.5	5.4	0.4
4	35	35	34	34.7	34-35	1	4	8.5	8.5	8.7	8.6	0.2
5	30	32	31	31.0	30-32	2	5	8.7	9.5	8.6	8.9	0.9
6	30	33	34	32.3	30-34	4	6	11.5	11.2	10.3	11.0	1.2
7	25	24	24	24.3	24-25	1	7	10.7	11.1	10.1	10.6	1.0
OVERALL	**217**	**220**	**224**	**220.3**	**217-223**	**6**	**OVERALL**	**8.8**	**8.6**	**8.5**	**8.6**	**0.3**

Probability Format: FREQUENCY

Word Count							Flesch-Kincaid Grade Level					
Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Range**	**Difference (Max-Min)**	Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Difference (Max-Min)**
1	25	27	26	26.0	25-27	2	1	10.0	9.7	9.5	9.7	0.5
2	28	25	28	27.0	25-28	3	2	12.0	11.1	12.0	11.7	0.9
3	30	29	33	30.7	29-33	4	3	5.2	5.6	5.5	5.4	0.4
4	35	35	34	34.7	34-35	1	4	8.5	8.5	8.7	8.6	0.2
5	31	33	32	32.0	31-33	2	5	8.9	9.7	8.8	9.1	0.9
6	31	34	35	33.3	31-35	4	6	10.8	10.6	9.8	10.4	1.0
7	26	25	25	25.3	25-26	1	7	11.3	11.7	10.7	11.2	1.0
OVERALL	**220**	**223**	**227**	**223.3**	**220-227**	**7**	**OVERALL**	**8.9**	**8.6**	**8.6**	**8.7**	**0.3**

Probability Format: WORDS

Word Count							Flesch-Kincaid Grade Level					
Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Range**	**Difference (Max-Min)**	Sub-section #	Med G	Proc A	Vaccine Q	**Mean**	**Difference (Max-Min)**
1	25	27	26	26.0	25-27	2	1	10.0	9.7	9.5	9.7	0.5
2	28	25	28	27.0	25-28	3	2	12.0	11.1	12.0	11.7	0.9
3	30	29	33	30.7	29-33	4	3	5.2	5.6	5.5	5.4	0.4
4	35	35	34	34.7	34-35	1	4	8.5	8.5	8.7	8.6	0.2
5	33	35	34	34.0	33-35	2	5	8.7	9.5	8.6	8.9	0.9
6	33	36	37	35.3	33-37	4	6	11.2	11.0	10.2	10.8	1.0
7	29	28	28	28.3	28-29	1	7	10.9	11.2	10.3	10.8	0.9
OVERALL	**227**	**230**	**234**	**230.3**	**227-234**	**7**	**OVERALL**	**8.9**	**8.7**	**8.6**	**8.7**	**0.3**

APPENDIX C

COUNTERBALANCE FOR PHASE 1

Table 4

Counterbalance for Phase 1

Order	Format-Passage 1	Format-Passage 2	Format-Passage 3
1	Percent-Vaccine Q	Frequency-Procedure A	Words-Medication G
2	Percent- Medication G	Frequency- Vaccine Q	Words - Procedure A
3	Percent- Procedure A	Frequency-Medication G	Words - Vaccine Q
4	Percent- Medication G	Frequency - Procedure A	Words - Vaccine Q
5	Percent- Procedure A	Frequency - Vaccine Q	Words - Medication G
6	Percent- Vaccine Q	Frequency-Medication G	Words - Procedure A
7	Words - Vaccine Q	Percent-Procedure A	Frequency-Medication G
8	Words - Medication G	Percent- Vaccine Q	Frequency - Procedure A
9	Words - Procedure A	Percent- Medication G	Frequency - Vaccine Q
10	Words - Medication G	Percent-Procedure A	Frequency - Vaccine Q
11	Words - Procedure A	Percent- Vaccine Q	Frequency-Medication G
12	Words - Vaccine Q	Percent-Medication G	Frequency - Procedure A
13	Frequency - Vaccine Q	Words - Procedure A	Percent- Medication G
14	Frequency-Medication G	Words - Vaccine Q	Percent- Procedure A
15	Frequency - Procedure A	Words - Medication G	Percent- Vaccine Q
16	Frequency-Medication G	Words - Procedure A	Percent- Vaccine Q
17	Frequency - Procedure A	Words - Vaccine Q	Percent- Medication G
18	Frequency - Vaccine Q	Words - Medication G	Percent- Procedure A
19	Words - Vaccine Q	Frequency - Procedure A	Percent- Medication G
20	Words - Medication G	Frequency - Vaccine Q	Percent- Procedure A
21	Words - Procedure A	Frequency-Medication G	Percent- Vaccine Q
22	Words - Medication G	Frequency - Procedure A	Percent- Vaccine Q
23	Words - Procedure A	Frequency - Vaccine Q	Percent- Medication G
24	Words - Vaccine Q	Frequency-Medication G	Percent- Procedure A

Order	Format-Passage 1	Format-Passage 2	Format-Passage 3
25	Frequency - Vaccine Q	Percent-P Procedure A	Words - Medication G
26	Frequency-Medication G	Percent- Vaccine Q	Words - Procedure A
27	Frequency - Procedure A	Percent- Medication G	Words - Vaccine Q
28	Frequency-Medication G	Percent- Procedure A	Words - Vaccine Q
29	Frequency - Procedure A	Percent- Vaccine Q	Words - Medication G
30	Frequency - Vaccine Q	Percent- Medication G	Words - Procedure A
31	Percent- Vaccine Q	Words - Procedure A	Frequency-Medication G
32	Percent- Medication G	Words - Vaccine Q	Frequency - Procedure A
33	Percent- Procedure A	Words - Medication G	Frequency - Vaccine Q
34	Percent- Medication G	Words - Procedure A	Frequency - Vaccine Q
35	Percent-Procedure A	Words - Vaccine Q	Frequency-Medication G
36	Percent- Vaccine Q	Words - Medication G	Frequency - Procedure A

APPENDIX D

VERBAL PHRASES OF PROBABILITIES

Verbal and Numeric Uncertainty Scales as They Appeared in Experiments 1 and 3

Verbal	Numeric (%)
__ Certain	__ 100
__ Almost totally certain	__ 95
__ Extremely likely	__ 90
__ Very likely	__ 85
__ Quite likely	__ 80
__ Likely	__ 75
__ Rather likely	__ 70
__ Fairly likely	__ 65
__ Somewhat likely	__ 60
__ Slightly likely	__ 55
__ As likely as is unlikely	__ 50
__ Slightly unlikely	__ 45
__ Somewhat unlikely	__ 40
__ Fairly unlikely	__ 35
__ Rather unlikely	__ 30
__ Unlikely	__ 25
__ Quite unlikely	__ 20
__ Very unlikely	__ 15
__ Extremely unlikely	__ 10
__ Almost totally impossible	__ 5
__ Impossible	__ 0

Figure 1. Verbal descriptions of probabilities taken from Windschitl & Wells (1996), p. 363.

APPENDIX E

MEDICATION G PASSAGE (FREQUENCY FORMAT)

Medication G Leaflet

What is Medication G?

Medication G is in a group of drugs called anti-inflammatory drugs. This medication works by reducing hormones that cause inflammation and pain in the body.

Medication G is used to reduce fever and treat pain or inflammation caused by many conditions such as headache, toothache, back pain, arthritis, menstrual cramps, or minor injury.

If a dose of Medication G is missed, take the missed dose as soon as remembered. However, skip the missed dose if it is almost time for the next dose.

What are the risks from Medication G?

Do not take more of this medication than is recommended. An overdose of Medication G can damage the stomach. Use only the smallest amount of medication needed to get relief from pain, swelling, or fever.

17 in 20 adults will experience mild problems after taking Medication G. These problems are noticeable but do not interfere with activities. An example of a mild problem is dry mouth.

1 in 5 adults will experience moderate problems after taking Medication G. These problems interfere with activities but do not require medical attention. An example of a moderate problem is nausea.

1 in 10,000 adults will experience severe problems after taking Medication G. These problems require medical attention. An example of a severe problem is blurred vision.

APPENDIX F

MEDICATION G PASSAGE (PERCENT FORMAT)

Medication G Leaflet

What is Medication G?

Medication G is in a group of drugs called anti-inflammatory drugs. This medication works by reducing hormones that cause inflammation and pain in the body.

Medication G is used to reduce fever and treat pain or inflammation caused by many conditions such as headache, toothache, back pain, arthritis, menstrual cramps, or minor injury.

If a dose of Medication G is missed, take the missed dose as soon as remembered. However, skip the missed dose if it is almost time for the next dose.

What are the risks from Medication G?

Do not take more of this medication than is recommended. An overdose of Medication G can damage the stomach. Use only the smallest amount of medication needed to get relief from pain, swelling, or fever.

85% of adults will experience mild problems after taking Medication G. These problems are noticeable but do not interfere with activities. An example of a mild problem is dry mouth.

20% of adults will experience moderate problems after taking Medication G. These problems interfere with activities but do not require medical attention. An example of a moderate problem is nausea.

0.01% of adults will experience severe problems after taking Medication G. These problems require medical attention. An example of a severe problem is blurred vision.

APPENDIX G

MEDICATION G PASSAGE (WORDS FORMAT)

Medication G Leaflet

What is Medication G?

Medication G is in a group of drugs called anti-inflammatory drugs. This medication works by reducing hormones that cause inflammation and pain in the body.

Medication G is used to reduce fever and treat pain or inflammation caused by many conditions such as headache, toothache, back pain, arthritis, menstrual cramps, or minor injury.

If a dose of Medication G is missed, take the missed dose as soon as remembered. However, skip the missed dose if it is almost time for the next dose.

What are the risks from Medication G?

Do not take more of this medication than is recommended. An overdose of Medication G can damage the stomach. Use only the smallest amount of medication needed to get relief from pain, swelling, or fever.

It is very likely that adults will experience mild problems after taking Medication G. These problems are noticeable but do not interfere with activities. An example of a mild problem is dry mouth.

It is quite unlikely that adults will experience moderate problems after taking Medication G. These problems interfere with activities but do not require medical attention. An example of a moderate problem is nausea.

It is almost totally impossible that adults will experience severe problems after taking Medication G. These problems require medical attention. An example of a severe problem is blurred vision.

APPENDIX H

PROCEDURE A PASSAGE (FREQUENCY FORMAT)

Procedure A Information Sheet

What is Procedure A?

The purpose of this medical procedure is to remove damaged tissue from the abdomen. This procedure will remove the infection that makes the abdomen painful and bloated.

The surgeon will make several small incisions in the skin on the abdomen through which a scope and hollow tubes called ports will be inserted.

The surgeon uses the scope to see inside the abdomen. Tools are passed through the ports inserted into the abdomen. The tools are used to remove the damaged tissue.

What are the risks from Procedure A?

Procedure A may not cure or relieve the condition or symptoms. The symptoms may come back and even worsen. Discuss treatment options and the various side effects associated with each option with a healthcare professional.

17 in 20 patients will experience mild problems after having Procedure A. These problems are noticeable but do not interfere with activities. An example of a mild problem is tenderness at incision sites.

1 in 5 patients will experience moderate problems after having Procedure A. These problems interfere with activities but do not require medical attention. An example of a moderate problem is swelling at incision sites.

1 in 10,000 patients will experience severe problems after having Procedure A. These problems require medical attention. An example of a severe problem is infection.

APPENDIX I

PROCEDURE A PASSAGE (PERCENT FORMAT)

Procedure A Information Sheet

What is Procedure A?

The purpose of this medical procedure is to remove damaged tissue from the abdomen. This procedure will remove the infection that makes the abdomen painful and bloated.

The surgeon will make several small incisions in the skin on the abdomen through which a scope and hollow tubes called ports will be inserted.

The surgeon uses the scope to see inside the abdomen. Tools are passed through the ports inserted into the abdomen. The tools are used to remove the damaged tissue.

What are the risks from Procedure A?

Procedure A may not cure or relieve the condition or symptoms. The symptoms may come back and even worsen. Discuss treatment options and the various side effects associated with each option with a healthcare professional.

85% of patients will experience mild problems after having Procedure A. These problems are noticeable but do not interfere with activities. An example of a mild problem is tenderness at incision sites.

20% of patients will experience moderate problems after having Procedure A. These problems interfere with activities but do not require medical attention. An example of a moderate problem is swelling at incision sites.

0.01% of patients will experience severe problems after having Procedure A. These problems require medical attention. An example of a severe problem is infection.

APPENDIX J

PROCEDURE A PASSAGE (WORDS FORMAT)

Procedure A Information Sheet

What is Procedure A?

The purpose of this medical procedure is to remove damaged tissue from the abdomen. This procedure will remove the infection that makes the abdomen painful and bloated.

The surgeon will make several small incisions in the skin on the abdomen through which a scope and hollow tubes called ports will be inserted.

The surgeon uses the scope to see inside the abdomen. Tools are passed through the ports inserted into the abdomen. The tools are used to remove the damaged tissue.

What are the risks from Procedure A?

Procedure A may not cure or relieve the condition or symptoms. The symptoms may come back and even worsen. Discuss treatment options and the various side effects associated with each option with a healthcare professional.

It is very likely that patients will experience mild problems after having Procedure A. These problems are noticeable but do not interfere with activities. An example of a mild problem is tenderness at incision sites.

It is quite unlikely that patients will experience moderate problems after having Procedure A. These problems interfere with activities but do not require medical attention. An example of a moderate problem is swelling at incision sites.

It is almost totally impossible that patients will experience severe problems after having Procedure A. These problems require medical attention. An example of a severe problem is infection.

APPENDIX K

VACCINE Q PASSAGE (FREQUENCY FORMAT)

Vaccine Q Information Statement

Why get vaccinated?

Infants and children are routinely vaccinated against Disease Q. But older children, adolescents, and adults need protection from this disease too. Vaccine Q provides that protection.

The United States averaged more than 100,000 cases of Disease Q each year before the vaccine. Since the vaccine has been available, Disease Q cases have fallen significantly.

Disease Q causes pain, a rash, a high fever, and it can be deadly. Disease Q is spread from person to person. Vaccine Q strengthens the body's ability to fight off Disease Q.

What are the risks from Vaccine Q?

With Vaccine Q, as with any medicine, there is always a risk of an allergic reaction. However, getting Disease Q would be much more likely to lead to severe problems than getting the vaccine.

17 in 20 adults will experience mild problems after getting Vaccine Q. These problems are noticeable but do not interfere with activities. An example of a mild problem is a low-grade fever.

1 in 5 adults will experience moderate problems after getting Vaccine Q. These problems interfere with activities but do not require medical attention. An example of a moderate problem is pain at the injection site.

1 in 10,000 adults will experience severe problems after getting Vaccine Q. These problems require medical attention. An example of a severe problem is seizures.

APPENDIX L

VACCINE Q PASSAGE (PERCENT FORMAT)

Vaccine Q Information Statement

Why get vaccinated?

Infants and children are routinely vaccinated against Disease Q. But older children, adolescents, and adults need protection from this disease too. Vaccine Q provides that protection.

The United States averaged more than 100,000 cases of Disease Q each year before the vaccine. Since the vaccine has been available, Disease Q cases have fallen significantly.

Disease Q causes pain, a rash, a high fever, and it can be deadly. Disease Q is spread from person to person. Vaccine Q strengthens the body's ability to fight off Disease Q.

What are the risks from Vaccine Q?

With Vaccine Q, as with any medicine, there is always a risk of an allergic reaction. However, getting Disease Q would be much more likely to lead to severe problems than getting the vaccine.

85% of adults will experience mild problems after getting Vaccine Q. These problems are noticeable but do not interfere with activities. An example of a mild problem is a low-grade fever.

20% of adults will experience moderate problems after getting Vaccine Q. These problems interfere with activities but do not require medical attention. An example of a moderate problem is pain at the injection site.

0.01% of adults will experience severe problems after getting Vaccine Q. These problems require medical attention. An example of a severe problem is seizures.

APPENDIX M

VACCINE A PASSAGE (WORDS FORMAT)

Vaccine Q Information Statement

Why get vaccinated?

Infants and children are routinely vaccinated against Disease Q. But older children, adolescents, and adults need protection from this disease too. Vaccine Q provides that protection.

The United States averaged more than 100,000 cases of Disease Q each year before the vaccine. Since the vaccine has been available, Disease Q cases have fallen significantly.

Disease Q causes pain, a rash, a high fever, and it can be deadly. Disease Q is spread from person to person. Vaccine Q strengthens the body's ability to fight off Disease Q.

What are the risks from Vaccine Q?

With Vaccine Q, as with any medicine, there is always a risk of an allergic reaction. However, getting Disease Q would be much more likely to lead to severe problems than getting the vaccine.

It is very likely that adults will experience mild problems after getting Vaccine Q. These problems are noticeable but do not interfere with activities. An example of a mild problem is a low-grade fever.

It is quite unlikely that adults will experience moderate problems after getting Vaccine Q. These problems interfere with activities but do not require medical attention. An example of a moderate problem is pain at the injection site.

It is almost totally impossible that adults will experience severe problems after getting Vaccine Q. These problems require medical attention. An example of a severe problem is seizures.

APPENDIX N

PRACTICE PASSAGE

The practice passage was an expository text not related to the health domain but comparable to the experimental passages in Flesch-Kincaid reading grade level at 8.4. The passage comprised four sections; each section had an average of 25 words (range: 20-30 words). The practice passage was an excerpt from Passage Two of the Nelson-Denny Reading Comprehension Test (Brown, Fischco, & Hanna, 1993).

Insects

Many insects communicate through sound. Male crickets use sound to attract females and to warn other males away from their territories.

Each cricket species produces several calls that differ from those of other cricket species. In fact, because many species look similar, entomologists often use the calls to identify the species.

Mosquitoes depend on sound, too. Males that are ready to mate home in on the buzzing sounds produced by females.

The male senses this buzzing by means of tiny hairs on his antennae, which vibrate only to the frequency emitted by a female of the same species.

APPENDIX O

PRACTICE COMPREHENSION QUESTIONS

1. Predators of mosquitoes would be very successful if they could detect female ________.

   A. wings
   B. buzzing
   C. singing
   D. antennae

2. Male crickets use sound to ____________.

   A. call other males
   B. frighten off females
   C. confuse their predators
   D. attract their mates

3. In the phrase "home in on the buzzing sounds," home means ____________.

   A. house
   B. listen
   C. focus
   D. join

4. Insects of the same species __________ to communicate with each other.

   A. have adapted
   B. sing
   C. speak
   D. have bonded

5. Scientists use insect ____________ to identify species.

   A. wings
   B. calls
   C. antennae
   D. legs

6. Male mosquitoes use the buzzing sound produced by females to _____________.

   A. locate food
   B. locate water
   C. identify a mate
   D. accompany their "songs"

APPENDIX P

MEDICATION G COMPREHENSION QUESTIONS

1. Medication G is used to treat ________.
    a. pain
    b. insomnia
    c. nausea
    d. vertigo

2. Upset stomach is _______________ with Medication G.
    a. treated
    b. a severe problem associated
    c. a moderate problem associated
    d. a mild problem associated

3. Moderate problems will be experienced by ________ of adults after taking Medication G.
    a. 1% -5%
    b. 50%-55%
    c. 20%-25%
    d. 75%-80%

4. If a dose of Medication G is missed, the missed dose should be ___________.
    a. skipped
    b. taken immediately
    c. taken only if the next dose is scheduled in more than one hour
    d. taken with the next dose

5. Medication G __________________ in the body that are responsible for causing discomfort.
    a. promotes the production of tissues
    b. decreases the production of agents
    c. decreases the production of fats
    d. increases the production of agents

6. A severe problem will be experienced by ____________ adults after taking Medication G.
    a. 1 in 1,000
    b. 1 in 100
    c. 1 in 10,000
    d. 1 in 10

7. It is ______________ that adults will experience moderate problems after taking Medication G.
    a. quite unlikely
    b. certain
    c. almost totally impossible
    d. very likely

8. It is __________that adults will experience blurred vision after taking Medication G.
    a. impossible
    b. unlikely
    c. almost totally impossible
    d. slightly unlikely

9. Medication G can_______________.
    a. increase body temperature
    b. lower blood pressure
    c. increase blood pressure
    d. lower body temperature

10. Medication G should be taken _____________ for the medication to work most effectively.
    a. every four hours
    b. with food
    c. sparingly
    d. before the pain returns

11. Medication G would not be used to treat __________.
    a. body aches
    b. allergies
    c. sore throat
    d. headache

12. _________ adults will experience mild problems after taking Medication G.
    a. 1 in 20
    b. 1 in 7
    c. 17 in 20
    d. 17 in 100

13. It is _____________ that adults will experience mild problems after taking Medication G.
    a. very likely
    b. almost totally impossible
    c. certain
    d. unlikely

14. _________ the usual dose of Medication G is recommended if a fever is higher than 102^{o}F.
    a. Twice
    b. Half
    c. Three times
    d. No change to

15. Approximately _______ adults will experience nausea after taking Medication G.
    a. 1 in 50
    b. 5 in 10
    c. 5 in 100
    d. 1 in 5

16. __________of adults will experience problems such as a dry mouth after taking Medication G.
    a. 20%-30%
    b. 80%-90%
    c. 1%-10%
    d. 50%-60%

17. Severe problems occur _________ mild problems.
    a. as often as
    b. less frequently than
    c. more frequently than
    d. twice as often as

18. Mild problems associated with Medication G include __________.
    a. bad breath
    b. difficulty chewing
    c. sore throat
    d. dry mouth

19. Too much Medication G can cause __________ damage.
    a. kidney
    b. liver
    c. stomach
    d. intestinal

20. Severe problems will be experienced by ________ of adults after taking Medication G.
    a. less than 50% but greater than 40%
    b. less than 20% but greater than 10%
    c. less than 10% but greater than 5%
    d. less than 1% but greater than 0%

APPENDIX Q

PROCEDURE A COMPREHENSION QUESTIONS

1. Moderate problems occur ___________ severe problems.
    a. as often as
    b. more frequently than
    c. less frequently than
    d. twice as often as

2. Procedure A removes the infection that makes the _________.
    a. stomach upset
    b. chest painful
    c. abdomen bloated
    d. kidneys malfunction

3. Sensitivity around stitches is an example of a ________ problem caused by Procedure A.
    a. rare
    b. severe
    c. moderate
    d. mild

4. Moderate problems will be experienced by ________ of patients after having Procedure A.
    a. 20%-25%
    b. 1% -5%
    c. 50%-55%
    d. 75%-80%

5. _______ patients will experience mild problems after having Procedure A.
    a. 17 in 100
    b. 1 in 20
    c. 17 in 20
    d. 1 in 7

6. It is __________ that adults will experience moderate problems after having Procedure A.
    a. very likely
    b. almost totally impossible
    c. certain
    d. quite unlikely

7. The source of pain will be removed through __________.
    a. hollow tubes
    b. scopes
    c. catheters
    d. stents

8. Severe problems will be experienced by _________ of patients after having Procedure A.
    a. less than 20% but greater than 10%
    b. less than 50% but greater than 40%
    c. less than 1% but greater than 0%
    d. less than 10% but greater than 5%

9. Procedure A will________ make the patient feel better.
    a. most likely
    b. always
    c. never
    d. almost certainly

10. Damaged tissue will be removed by having a ______________.
    a. large incision on the abdomen
    b. scope put down the throat and passed down to the stomach
    c. medicine injected into the damaged tissue
    d. thin tube put into the abdomen through a small incision

11. It is ______ that patients will experience mild problems after having Procedure A.
    a. almost totally impossible
    b. very likely
    c. unlikely
    d. certain

12. It is __________that patients will experience infection after having Procedure A.
    a. almost totally impossible
    b. impossible
    c. unlikely
    d. slightly unlikely

13. Approximately ________ patients will experience swelling at incision sites after having Procedure A.
    a. 5 in 10
    b. 5 in 100
    c. 1 in 5
    d. 1 in 50

14. _________________, which should be discussed with a physician.
    a. All treatment options are without risk
    b. All treatment options have deadly side effects
    c. Only Procedure A has minimal risks
    d. Each treatment option has risks

15. A severe problem is experienced by ____________ patients after having Procedure A.
    a. 1 in 10
    b. 1 in 100
    c. 1 in 1,000
    d. 1 in 10,000

16. Procedure A may not cure the condition or _________ symptoms.
    a. relieve
    b. cause
    c. cure
    d. treat

17. The tools and scope used for Procedure A are ________.
    a. long and narrow
    b. short and narrow
    c. long and wide
    d. short and wide

18. Patients who have Procedure A will have __________ scar(s).
    a. many large
    b. one small
    c. one large
    d. many small

19. __________of patients will experience problems such as tenderness at incision sites after having Procedure A.
    a. 1%-10%
    b. 20%-30%
    c. 50%-60%
    d. 80%-90%

20. An example of a severe problem experienced by patients after having Procedure A is ___________.
    a. liver damage
    b. infection
    c. excessive weight gain
    d. excessive weight loss

APPENDIX R

VACCINE Q COMPREHENSION QUESTIONS

1. _______________ adults will experience mild problems after getting Vaccine Q.
    a. 1 in 20
    b. 17 in 20
    c. 17 in 100
    d. 1 in 7

2. There is ______________ a risk of a having a serious problem after getting Vaccine Q, as there is with any medicine.
    a. never
    b. sometimes
    c. often
    d. always

3. Disease Q was a public health _____________ before the vaccine.
    a. loss
    b. threat
    c. topic
    d. question

4. Approximately _________ adults will experience pain at the injection site after getting Vaccine Q.
    a. 1 in 5
    b. 1 in 50
    c. 5 in 10
    d. 5 in 100

5. Adults who get Disease Q might experience ____________________.
    a. a temperature of 104^{o}F
    b. dizziness
    c. difficulty breathing
    d. acute coughing

6. Vaccine Q _______________________ Disease Q.
    a. increases the symptoms of
    b. protects against
    c. causes
    d. reduces the symptoms of

7. It is __________________ that adults will experience moderate problems after getting Vaccine Q.
    a. very likely
    b. quite unlikely
    c. certain
    d. almost totally impossible

8. Deaths from Disease Q have_________________ after Americans began getting Vaccine Q regularly.
    a. increased
    b. decreased dramatically
    c. not changed
    d. decreased slightly

9. Severe problems will be experienced by _____________ of adults after getting Vaccine Q.
    a. less than 1% but greater than 0%
    b. less than 10% but greater than 5%
    c. less than 20% but greater than 10%
    d. less than 50% but greater than 40%

10. Moderate problems will be experienced by ___________ of adults after getting Vaccine Q.
    a. 50%-55%
    b. 75%-80%
    c. 20%-25%
    d. 1% -5%

11. Convulsions are an example of a ___________ problem associated with getting Vaccine Q.
    a. frequent
    b. mild
    c. severe
    d. moderate

12. Because Disease Q is _______________, it is recommended to restrict contact with others until there have been no symptoms for at least 24 hours.
    a. non-communicable
    b. harmless
    c. contagious
    d. benign

13. It is ________________ that adults will experience seizures after getting Vaccine Q.
    a. slightly unlikely
    b. impossible
    c. unlikely
    d. almost totally impossible

14. _______________of adults experience problems such as a low-grade fever after getting Vaccine Q.
    a. 80%-90%
    b. 20%-30%
    c. 1%-10%
    d. 50%-60%

15. Getting Vaccine Q will ______________________ Disease Q.
    a. do nothing to protect against
    b. boost the body's protection against
    c. increase the chance of getting Disease Q
    d. reduce the body's protection against

16. A severe problem will be experienced by ____________ adults after getting Vaccine Q.
    a. 1 in 10,000
    b. 1 in 1,000
    c. 1 in 100
    d. 1 in 10

17. Mild problems occur ___________________ moderate problems.
    a. more frequently than
    b. as often as
    c. less frequently than
    d. twice as often as

18. An example of a moderate problem experienced by adults after getting Vaccine Q is _____________.
    a. body aches
    b. high fever
    c. rash at injection site
    d. pain at injection site

19. It is ________________ that adults will experience mild problems after getting Vaccine Q.
    a. certain
    b. very likely
    c. unlikely
    d. almost totally impossible

20. Getting Disease Q is _____________ than getting Vaccine Q.
    a. less expensive
    b. less dangerous
    c. more expensive
    d. more dangerous

APPENDIX S

ANSWERS FOR COMPREHENSION QUESTIONS BY CONTENT AND TYPE

MEDICATION G			
Content	**Type**	Question	Answer
General	Explicit	1	a
General	Inference	2	c
Probability	Explicit	3	c
General	Inference	4	c
General	Inference	5	b
Probability	Inference	6	c
Probability	Inference	7	a
Probability	Inference	8	c
General	Inference	9	d
General	Inference	10	c
General	Inference	11	b
Probability	Inference	12	c
Probability	Inference	13	a
General	Inference	14	d
Probability	Inference	15	d
Probability	Explicit	16	b
Probability	Inference	17	b
General	Explicit	18	d
General	Explicit	19	c
Probability	Explicit	20	d

PROCEDURE A			
Content	**Type**	Question	Answer
Probability	Inference	1	b
General	Explicit	2	c
General	Inference	3	d
Probability	Explicit	4	a
Probability	Inference	5	c
Probability	Inference	6	d
General	Inference	7	a
Probability	Explicit	8	c
General	Inference	9	a
General	Inference	10	d
Probability	Inference	11	b
Probability	Inference	12	a
Probability	Inference	13	c
General	Inference	14	d
Probability	Inference	15	d
General	Explicit	16	a
General	Inference	17	a
General	Inference	18	d
Probability	Explicit	19	d
General	Explicit	20	b

VACCINE Q			
Content	**Type**	Question	Answer
Probability	Inference	1	b
General	Explicit	2	d
General	Inference	3	b
Probability	Inference	4	a
General	Inference	5	a
General	Explicit	6	b
Probability	Inference	7	b
General	Inference	8	b
Probability	Explicit	9	a
Probability	Explicit	10	c
General	Inference	11	c
General	Inference	12	c
Probability	Inference	13	d
Probability	Explicit	14	a
General	Inference	15	b
Probability	Inference	16	a
Probability	Inference	17	a
General	Explicit	18	d
Probability	Inference	19	b
General	Inference	20	d

APPENDIX T

FORMAT COUNTERBALANCE FOR MAGNITUDE COMPARISON TASK

Table 1

Format Counterbalance for Magnitude Comparison Task

Order	**Format 1**	**Format 2**	**Format 3**
1	Percent	Frequency	Words
2	Words	Percent	Frequency
3	Frequency	Words	Percent
4	Words	Frequency	Percent
5	Frequency	Percent	Words
6	Percent	Words	Frequency

REFERENCES

Adams, C., Smith, M. C., Pasupathi, M., & Vitolo, L. (2002). Social context effects on story recall in older and younger women: Does the listener make a difference? *The Journals of Gerontology, 57B,* 28-40.

Apter, A. J., Cheng, J., Small, D., Bennett, I. M., Albert, C., Fein, D. G., George, M., & Van Horne, S. (2006). Asthma numeracy skill and health literacy. *The Journal of Asthma: Official Journal of the Association for the Care of Asthma, 43*, 705-710.

Arozullah, A. M., Yarnold, P. R., Bennett, C. L., Soltysik, R. C., Wolf, M. S., Ferreira, R. M., Lee, D. S., Costello, S., Shakir, A., Denwood, C., Bryant, F. B., Davis, T. (2007). Development and validation of a short-form, Rapid Estimate of Adult Literacy in Medicine. *Medical Care, 45,* 1026-1033.

Berry, D. C., Knapp, P., Raynor, D. K., (2002). Provision of information about drug side effects to patients. *Lancet, 359,* 853-854.

Bonato, M., Fabbri, S., Umilta , C., & Zorzi, M. (2007). The mental representation of numerical fractions: Real or integer? *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1410-1419.

Brown, J.I., Fishco, V.V. & Hanna, G. (1993). The Nelson Denny reading test. Chicago, IL: The Riverside Publishing Company.

Burkell, J. B. (2004). What are the chances? Evaluating risk and benefit information in consumer health materials. *Journal of the Medical Library Association, 92,* 200-208.

Calman, K. C. (1996). Cancer: Science and society and the communication of risk. *British Medical Journal, 313,* 799-802.

Cavanaugh, K., Huizinga, M. M., Wallston, K. A., Gebretsadik, T., Shintani, A., Davis, D., Gregory, R. P., Fuchs, L. Malone, R., Cherrington, A., Pignone, M., DeWalt, D. A. Elasy, T. A., & Rothman, R. L. (2008). Association of numeracy and diabetes control. *Annals of Internal Medicine, 148*, 737-746.

Centers for Disease Control and Prevention (CDC) Tetanus, Diphtheria (Td) or Tetanus, Diphtheria, Pertussis (Tdap) vaccine information statement. Retrieved May 28, 2010, http://www.cdc.gov/vaccines/pubs/vis/default.htm#tdtdap

Couper, M. P., & Singer, E. (2009). The role of numeracy in informed consent for surveys. *Journal of Empirical Research on Human Research Ethics, 4*, 17-26. doi:10.1525/jer.2009.4.4.17

Cuite, C. L., Weinstein, N. D., Emmons, K., & Colditz, G. (2008). A test of numeric formats for communicating risk probabilities. *Medical Decision Making, 28,* 377-384.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371-396.

Denes-Raj, V., Epstein, S., & Cole, J. (1995). The generality of the ratio-bias phenomenon. *Personality and Social Psychology Bulletin*, *21*, 1083-1092. doi:10.1177/01461672952110009

DeWolf, M. & Vosniadou, S. (2011). The whole number bias in fraction magnitude comparisons with adults. *CogSci2011 Proceedings*, 1751-1756.

Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis, 29,* 1473-1488.

Doak, C. C., Doak, L. G., Root, J. H. (1996). *Teaching patients with low literacy skills* (2nd ed.). Philadelphia, PA: J.B. Lippincott Co.

Durso, F. T., Rawson, K. A., & Girotto, S. (2007). Comprehension and situation awareness. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, T. J. Perfect (Eds.). *Handbook of applied cognition (*2nd ed., pp. 164-193). Hoboken, NJ US: John Wiley & Sons Inc.

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Services.

E-Prime 2. Psychological Software Tools, 2003

Estrada, C. A., Martin-Hryniewicz, M., Peek, B. T., Collins, C., & Byrd, J. C. (2004). Literacy and numeracy skills and anticoagulation control. *The American Journal of the Medical Sciences, 328,* 88-93.

Fagerlin, A., Ubel, P. A., Smith, D. M., & Zikmund-Fisher, B. J. (2007). Making numbers matter: Present and future research in risk communication. *American Journal of Health Behavior, 31(Suppl1),* S47-S56.

Farrell, M. H., Kuruvilla, P., Eskra, K. L., Christopher, S. A., & Brienza, R. S. (2009). A method to quantify and compare clinicians' assessments of patient understanding during counseling of standardized patients. *Patient Education and Counseling, 77,* 128-135.

Fausset, C. B. (2009). A review and analysis of health numeracy in younger and older adults: Toward a model of health numeracy. *Unpublished manuscript.*

Fink, A. S., Prochazka, A. V., Henderson, W. G., Bartenfeld, D., Nyirenda, C., Webb, A., Berger, D. H., Itani, K., Whitehill, T., Edwards, J., Wilson, M., Karsonovich, C., & Parmelee, P. (2010). Enhancement of surgical informed consent by addition of repeat back: a multicenter, randomized controlled clinical trial. *Annals of Surgery*, 252(1), 27-36.

Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging*, *25*, 271-288. doi:10.1037/a0019106

Finucane, M. L., Mertz, C. K., Slovic, P., & Schmidt, E. (2005). Task complexity and older adults' decision-making competence. *Psychology and Aging*, *20*, 71-84. doi:10.1037/0882-7974.20.1.71

Finucane, M. L., Slovic, P., Hibbard, J. H., Peters, E., Mertz, C. K., & MacGregor, D. G. (2002). Aging and decision-making competence: An analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *Journal of Behavioral Decision Making, 15*, 141-164.

Fuller, R., Dudley, N., & Blacktop, J. (2001). Risk communication and older people-understanding of probability and risk information by medical inpatients aged 75 years and older. *Age and Ageing, 30*, 473-476.

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

Galesic, M., Gigerenzer, G., & Straubinger, N. (2009). Natural frequencies help older adults and people with low numeracy to evaluate medical screening tests. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *29*, 368-371.

Gallistel & Gelman. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Science*, *4*, 59-65.

Geary, D. C., & Lin, J. (1998). Numerical cognition: Age-related differences in the speed of executing biologically primary and biologically secondary processes. *Experimental Aging Research, 24,* 101-137.

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risk: From innumeracy to insight. *British Medical Journal, 327,* 741-744.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53-96. doi:10.1111/j.1539-6053.2008.00033.x

Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (2005). 'A 30% chance of rain tomorrow': How does the public understand probabilistic weather forecasts? *Risk Analysis*, *25*, 623-629. doi:10.1111/j.1539-6924.2005.00608.x

Ginde, A. A., Clark, S., Goldstein, J. N., & Camargo, C. A., Jr. (2008). Demographic disparities in numeracy among emergency department patients: evidence from two multicenter studies. *Patient Education and Counseling, 72*, 350-356.

Gramling, R., Irvin, J., Nash, J., Sciamanna, C., & Culpepper, L. (2004). Numeracy and medicine: Key family physician attitudes about communicating probability with patients. *The Journal of the American Board of Family Practice, 17*, 473.

Halberda, J., Mazzocco, M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665–668.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 163-169. doi: 10.1037/0278-7393.17.1.163

Health Literacy Measurement Tools. January 2009. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/populations/sahlsatool.htm

Henneman, L., Marteau, T., & Timmermans, D. (2008). Clinical geneticists' and genetic counselors' views on the communication of genetic risks: a qualitative study. *Patient Education and Counseling*, *73*, 42-49.

Hoyer, W. J., & Verhaegen, P. (2006). Memory aging. In J. E. Birren & K. W. Schaie (Eds.) *Handbook of psychology of aging* (6^{th} ed., pp. 209-232). Amsterdam: Elsevier.

Hyman, I. (1994). Conversational remembering: Story recall with a peer versus for an experimenter. *Applied Cognitive Psychology, 8,* 49-66.

Johnson, R. E. (2003). Aging and the remembering of text. *Developmental Review, 23,* 261-346.

Karelitz, T. M., & Budescu, D. V. (2004). You say 'probable' and I say 'likely': Improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology: Applied*, *10*, 25-41. doi:10.1037/1076-898X.10.1.25

Knapp, P., Gardner, P. H., Carrigan, N., Raynor, D. K., & Woolf, E. (2009). Perceived risk of medicine side effects in users of a patient information website: A study of the use of verbal descriptors, percentages and natural frequencies. *British Journal of Health Psychology, 14,* 579-594.

Knapp, P., Raynor, D. K., & Berry, D. C. (2004). Comparison of two method of presenting risk information to patients about the side effects of medicines. *Quality and Safety in Health Care, 13*, 176-180.

Kutner, M., Greenberg, E., & Baer, C. (2005). A first look at the literacy of America's adults in the 21st century. National Center for Education Statistics Publication No. 2006-470. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

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.

Lipkus, I. M., & Peters, E. (2009). Understanding the role of numeracy in health: Proposed theoretical framework and practical insights. *Health Education and Behavior, 36*, 1065-1081.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making: An International Journal of the Society for Medical Decision Making, 21*, 37-44.

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: An International Journal of the Society for Medical Decision Making, 21,* 141-149.

Mazur, D., & Hickam, D. (1991). Patients' interpretations of probability terms. *Journal of General Internal Medicine*, *6*, 237-240.

Meyer, B. J. F., & Pollard, C. K (2006). Applied learning and aging: A closer look at reading. In J. E. Birren & K. W. Schaie (Eds.) *Handbook of psychology of aging* (6th ed., pp. 233-260). Amsterdam: Elsevier.

Morrow, D. G., & Rogers, W. A. (2008). Environmental support: An integrative framework. *Human Factors, 50*, 589-613.

Morrow, D. G., Weiner, M., Young, J., Steinley, D., Deer, M., & Murray, M. D. (2005). Improving medication knowledge among older adults with heart failure: A patient-centered approach to instruction design. *The Gerontologist, 45,* 545-552.

Mosteller, F. & Youtz, C. (1990). Quantifying probabilistic expressions. *Statistical Science, 5*, 2-34.

Moyer, R. S. & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519-1520. doi:10.1038/2151519a0

Natter, H. M., & Berry, D. C. (2005). Effects of active information processing on the understanding of risk information. *Applied Cognitive Psychology*, *19*, 123-135. doi:10.1002/acp.1068

Nees, M. A. & Walker, B. N. (2011). Mental scanning of sonifications reveals flexible encoding of nonspeech sounds and a universal per-item scanning cost. *Acta Psychologica, 137*, 309-317.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine, 35*, 261-274.

Ni, Y. & Zhou, Y. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40,* 27-52.

Nurss, J. R., Parker, R. M., Williams, M. V., & Baker, D. W. (2001). Test of functional health literacy in adults (TOFHLA) (2nd ed.). Snow Camp, NC: Peppercorn Books & Press, Inc.

Office of Disease Prevention and Health Promotion. Quick guide to health literacy: fact sheets, strategies, and resources. http://www.health.gov/communication/literacy/quickguide/Quickguide.pdf. Accessed July 7, 2010.

Pak, R., Czaja, S. J., Sharit, J., Rogers, W. A., & Fisk, A. D. (2008). The role of spatial abilities and age in performance in an auditory computer navigation task. *Computers in Behavior, 24, 3045-3051.*

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17,* 299-320.

Paulos, J. A. (1989). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Hill and Wang.

Peters, E., Slovic, P., Västfjäll, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment and Decision Making*, *3*, 619-635.

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

Rapid Estimate of Adult Literacy in Medicine short form (REALM-SF). Retrieved April 14, 2011 from http://www.ahrq.gov/populations/sahlsatool.htm

Reyna, V. F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making: An International Journal of the Society for Medical Decision Making, 28,* 850-865. doi:10.1177/0272989X08327066

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, *17,* 147-159. doi:10.1016/j.lindif.2007.03.010

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

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

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle longitudinal study*. Cambridge: Cambridge University Press.

Schaie, K. W. (2007). Generational differences: The age-cohort period model. In J. E. Birren (Ed.), *Encyclopedia of Gerontology: Age, aging and the Aged* (2nd ed., pp. 601-609). Oxford, UK: Elsevier.

Schapira, M. M., Fletcher, K. E., Gilligan, M. A., King, T. K., Laud, P. W., Matthews, B. A., & Neuner, J. M. (2008). A framework for health numeracy How patients use quantitative skills in health care. *Journal of Health Communication, 13,* 501-517.

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,* 459-467.

Schillinger, D., Piette, J., Grumbach, K., Wang, F., Wilson, C., Daher, C., Leong-Grotz, K., Castro, C., Bindman, A. B. (2003). Closing the loop: Physician communication with diabetic patients who have low health literacy. *Archives of Internal Medicine, 163,* 83-90.

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1227-1238.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*, 966-972.

Schwartz, L., Woloshin, S., & Welch, H. (2005). Can patients interpret health information? An assessment of the medical data interpretation test. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *25,* 290-300.

Sheridan, S., & Pignone, M. (2002). Numeracy and the medical student's ability to interpret data. *Effective Clinical Practice, 5,* 35-40.

Shipley, W. C. (1986). Shipley Institute of Living Scale. Los Angeles: Western Psychological Services.

Tudusciuc, O., & Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proceedings of the National Academy of Sciences of the USA*, *104*, 14513-14518.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and non-linear age effects and structural models. *Psychological Bulletin*, *122*, 231-249.

Waldrop-Valverde, D., Jones, D. L., Jayaweera, D., Gonzalez, P., Romero, J., & Ownby, R. L. (2009). Gender differences in medication management capacity in HIV infection: The role of health literacy and numeracy. *AIDS and Behavior*, *13*, 46-52.

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General*, *115*, 348-365. doi:10.1037/0096-3445.115.4.348

Wechsler, D. (1997). Wechsler Adult Intelligence Scale III. (3rd ed.). San Antonio, TX: The Psychological Corporation.

Weinstein, N. (1999). What does it mean to understand a risk? Evaluating risk comprehension. *Journal of the National Cancer Institute Monographs*, *25*, 15-20.

Wills, C. E., Holmes-Rovner, M. (2003). Patient comprehension of information for shared treatment decision making: State of the art and future directions. *Patient Education & Counseling*, *50*, 285–290.

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied*, *2*, 343-364. doi:10.1037/1076-898X.2.4.343

Woloshin, K., Ruffin, M., & Gorenflo, D. (1994). Patients' interpretation of qualitative probability statements. *Archives of Family Medicine, 3*, 961-966.

Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology, 11*, 495-506.